May 2019

Tornado Hazard Analysis for the United States using a Stochastic Track Simulation Model

Fanfu Fan
Clemson University, fanfanfu2011@gmail.com

Follow this and additional works at: https://tigerprints.clemson.edu/all_dissertations

Recommended Citation

This Dissertation is brought to you for free and open access by the Dissertations at TigerPrints. It has been accepted for inclusion in All Dissertations by an authorized administrator of TigerPrints. For more information, please contact kokeefe@clemson.edu.
Tornado Hazard Analysis for the United States using a Stochastic Track Simulation Model

A Dissertation
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
Civil Engineering

by
Fanfu Fan
May 2019

Accepted by:
Weichiang Pang, Committee Chair
Thomas E. Cousins
Brandon Ross
Nadarajah Ravichandran
ABSTRACT

During the period from 1950 to 2015, the United States experienced more than 60,000 tornadoes resulting in more than 900,000 injuries and about 6,000 fatalities (NOAA, 2016). Compared to hurricanes, the impact of a tornado is much localized and the probability of occurrence at a given location can be extremely low. Therefore, it is not feasible to use solely the raw historical data or tracks to quantify the risk of tornadoes for a given structure or a city that has not been affected by historical tornadoes. In order to properly quantify the risk of tornado, there is a need to develop a stochastic tornado simulation model to generate a large database of synthetic tornado tracks to quantify the tornado hazard. To carry out tornado risk assessment, both a methodology to perform stochastic simulation of tornado tracks and a tornado risk analysis framework are needed for the continental United States and the details of these frameworks will be presented in the following study.

In Chapter 2, a methodology to perform stochastic simulation of tornado tracks and parameters is presented. The stochastic simulation framework contains three sub-models, namely, genesis model, tracking model and wind field model. The genesis model utilizes the kernel density estimation method to simulate the annual number of tornadoes and touchdown locations. The tracking model is utilized to generate the tornado intensity, path width, path length, heading direction, intensity and time/date of spawn. The wind field model was used to compute the tornado wind speed along the tornado footprint. The tracking model was calibrated using the historical tornado information maintained by the NOAA Storm Prediction Center (SPC). A database of 1 million years of simulated tornado
tracks was generated using the Clemson high performance computing facility. The final simulated tornado track parameters include the tornado occurrence rate, intensity (EF scale), touchdown location, touchdown time, and path direction. All these parameters are geographic dependent, in other words, the simulated parameters vary spatially and depending on its spawn locations.

Chapter 3 presents a framework of develop the tornado hazard maps in United States. Using the simulated tornado database (Chapter 2), Hazard maps in United States for EF0-EF5 wind speeds have been developed for several different target structure sizes, and the target include point target, 0.08 mi$^2$, 0.03 mi$^2$, and 0.5 mi$^2$ circular target, respectively. Relationship between tornado striking probability and target size have been investigated, and tornado hazard for a specific structure in United States can be interpolated from given location and size using the hazard maps.

In order to predict the tornado damage and improve the community resilience performance, a new approach of tornado scenario selection and damage estimation is proposed in Chapter 4. The damage area and peak wind speed have been calculated, for each tornado tracks which impact the study domain, to estimate the corresponding mean recurrence interval (MRI). The building locations and dimensions are determined using an image segmentation algorithm, and the damage state is evaluated using the fragility curves. Damage estimation for three tornado scenarios, selected according to damage area, peak wind speed and both intensity measures were conducted with different hazard level (MRI).
DEDICATION

I dedicate this study to my parents, Surong Guo and Yanmin Fan for their love and support; and my beloved wife, Guangwen Mou.
ACKNOWLEDGMENTS

I would foremost like to thank my advisor Dr. Weichiang Pang for giving me such opportunity to work on this project in Clemson University. I would also like to thank my PhD committee, Dr. Thomas Cousins, Dr. Brandon Ross and Dr. Nadarajah Ravichandran for their comments and help. In addition, I would like to give my appreciations to all the teachers and stuffs of the Glenn Department of Civil Engineering, this work would not have been possible without their teaching, help and advice.

Last, but not least, I would like to thank my friends and colleagues, Xiaoyu Hu, Yuting Cheng, Jin Wang, Zhengshou Lai, Nathan Schneider, Michael Stoner, and Christopher Cornett, who were always willing to help and making my PhD life so enjoyable.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>TITLE PAGE</td>
<td>i</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>DEDICATION</td>
<td>iv</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>v</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>x</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xi</td>
</tr>
<tr>
<td>CHAPTER 1 INTRODUCTION</td>
<td>15</td>
</tr>
<tr>
<td>1.1 Tornado Basics and Background</td>
<td>15</td>
</tr>
<tr>
<td>1.1.1 Tornado spawn location</td>
<td>15</td>
</tr>
<tr>
<td>1.1.2 Tornado spawn time and season</td>
<td>17</td>
</tr>
<tr>
<td>1.1.3 Tornado path direction</td>
<td>19</td>
</tr>
<tr>
<td>1.1.4 Tornado magnitude</td>
<td>20</td>
</tr>
<tr>
<td>1.1.5 Tornado path length and path width</td>
<td>21</td>
</tr>
<tr>
<td>1.1.6 Tornado caused losses and fatalities</td>
<td>21</td>
</tr>
<tr>
<td>1.1.7 Tornado database</td>
<td>22</td>
</tr>
<tr>
<td>1.1.8 Methodology</td>
<td>23</td>
</tr>
<tr>
<td>1.2 Objectives</td>
<td>24</td>
</tr>
<tr>
<td>1.3 References</td>
<td>26</td>
</tr>
<tr>
<td>CHAPTER 2 TORNADO SIMULATION MODELS</td>
<td>29</td>
</tr>
<tr>
<td>2.1 Abstract</td>
<td>29</td>
</tr>
<tr>
<td>2.2 Introduction</td>
<td>30</td>
</tr>
</tbody>
</table>
2.3 Method of analysis ........................................................................................................ 34
  2.3.1 Data sources description and processing .............................................................. 34
  2.3.2 Stochastic Track Simulation Model .................................................................. 36
    2.3.2.1 Genesis Model .......................................................................................... 38
    2.3.2.2 Tracking model .......................................................................................... 43
    2.3.2.3 Wind field model ...................................................................................... 54
      2.3.2.3.1 Wind field along the tornado length .................................................. 54
      2.3.2.3.2 Wind field along the tornado width .................................................. 55
  2.4 Stochastic Tornado Track Simulation Program .................................................. 57
  2.5 Synthetic Tornado Tracks Database and Applications ........................................ 58
    2.5.1 Tornado tracks ............................................................................................... 59
    2.5.2 Tornado intensity ........................................................................................... 61
    2.5.3 Tornado spawn month .................................................................................. 63
    2.5.4 Tornado spawn hour in a day ......................................................................... 64
    2.5.5 Comparisons for Select Cities ....................................................................... 66
      2.5.5.1 Tornado spawn month/time ................................................................. 68
      2.5.5.2 Tornado Intensity ................................................................................. 69
      2.5.5.3 Hazard curves ...................................................................................... 71
  2.6 Summary and Conclusion ..................................................................................... 72
  2.7 References ............................................................................................................. 73

CHAPTER 3 Tornado Hazard Analysis and Effect of Structure Size............................ 76
  3.1 Abstract ................................................................................................................. 76
  3.2 Introduction ........................................................................................................... 77
  3.3 Tornado track database ....................................................................................... 78
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1.1: Tornado intensity and its corresponding wind speed.</td>
<td>20</td>
</tr>
<tr>
<td>Table 2.1: Tornado parameter groups.</td>
<td>44</td>
</tr>
<tr>
<td>Table 2.2: Modeled Weibull distribution parameters for path length and path width by EF scale.</td>
<td>50</td>
</tr>
<tr>
<td>Table 2.3: Tornado length group, ( L_i )</td>
<td>52</td>
</tr>
<tr>
<td>Table 2.4: Tornado width group, ( W_i )</td>
<td>52</td>
</tr>
<tr>
<td>Table 2.5: Number of simulated and observed tornado tracks near Oklahoma City</td>
<td>59</td>
</tr>
<tr>
<td>Table 3.1: Target size information</td>
<td>85</td>
</tr>
<tr>
<td>Table 3.2: Target size applied in hazard maps</td>
<td>88</td>
</tr>
<tr>
<td>Table 4.1: Correlation coefficients of tornado damage area and peak wind speed.</td>
<td>120</td>
</tr>
<tr>
<td>Table 4.2: Numbers of tornados scenarios selected for a given MRI range.</td>
<td>124</td>
</tr>
<tr>
<td>Table 4.3: Building Types and Description. (Adapted from Memari et al 2018.)</td>
<td>128</td>
</tr>
<tr>
<td>Table 4.4: Numbers of tornadoes induced building damages.</td>
<td>132</td>
</tr>
<tr>
<td>Table 4.5: Average numbers of damaged buildings caused by EF0 tornado.</td>
<td>134</td>
</tr>
<tr>
<td>Table 4.6: Average numbers of damaged buildings caused by EF1 tornado.</td>
<td>134</td>
</tr>
<tr>
<td>Table 4.7: Average numbers of damaged buildings caused by EF2 tornado.</td>
<td>134</td>
</tr>
<tr>
<td>Table 4.8: Average numbers of damaged buildings caused by EF3 tornado.</td>
<td>135</td>
</tr>
<tr>
<td>Table 4.9: Average numbers of damaged buildings caused by EF4 tornado.</td>
<td>135</td>
</tr>
<tr>
<td>Table 4.10: Average numbers of damaged buildings caused by EF5 tornado.</td>
<td>135</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.1: Traditional tornado alley, adopted from NOAA (2016)</td>
<td>16</td>
</tr>
<tr>
<td>Figure 1.2: Return frequency of F2+ tornadoes, 1961-2010 (adopted from Carbin et al. 2013)</td>
<td>17</td>
</tr>
<tr>
<td>Figure 1.3: Tornado averages by month 1991-2010 (adopted from NOAA, 2016)</td>
<td>18</td>
</tr>
<tr>
<td>Figure 1.4: Tornado time of occurrence (NOAA, 2016)</td>
<td>19</td>
</tr>
<tr>
<td>Figure 2.1: Annual number of tornadoes recorded from 1950 to 2015 and 10-year moving average</td>
<td>31</td>
</tr>
<tr>
<td>Figure 2.2: Cumulative Density Function (CDF) of tornado annually spawn frequency</td>
<td>38</td>
</tr>
<tr>
<td>Figure 2.3: Quantile-to-Quantile plot of observed versus modeled annual tornado spawn frequencies</td>
<td>39</td>
</tr>
<tr>
<td>Figure 2.4: Modeled tornado location by using multivariate normal random number generator</td>
<td>40</td>
</tr>
<tr>
<td>Figure 2.5: Probability density contours of tornado spawn locations for (A) modeled, and (B) observed tornadoes using the tornadoes spawned in year 1992 as the seeds</td>
<td>42</td>
</tr>
<tr>
<td>Figure 2.6: Probability density contours for (A) EF 2 tornadoes and (B) tornadoes spawned in early January</td>
<td>45</td>
</tr>
<tr>
<td>Figure 2.7: Probability density function of the contiguous United States tornadoes by EF scale</td>
<td>46</td>
</tr>
<tr>
<td>Figure 2.8: Site specific probability density function for a location in Oklahoma City by EF scale</td>
<td>47</td>
</tr>
<tr>
<td>Figure 2.9: Probability density function of the contiguous United States tornadoes by spawn hour</td>
<td>48</td>
</tr>
<tr>
<td>Figure 2.10: Probability density function of the contiguous United States tornadoes by spawn month</td>
<td>48</td>
</tr>
</tbody>
</table>
Figure 2.11: Cumulative probability distribution of tornado spawn day of the year
(day 1 to 365)............................................................................................................. 49
Figure 2.12: Cumulative density function (CDF) of tornado path length............... 50
Figure 2.13: Cumulative density function (CDF) of tornado path width. ............... 51
Figure 2.14: Small-scale tornado spawn location density contours for (A) length
less than 0.33 km, and (B) width less than 14.4 m. ................................................. 53
Figure 2.15: Large-scale tornado spawn location density contours for (A) length
greater than 5.5 km, and (B) width greater than 112.9 m. ...................................... 54
Figure 2.16: Tornado intensity variation along the track......................................... 55
Figure 2.17: Flowchart for stochastic tornado track simulation program............... 57
Figure 2.18: (A) Actual observed tornado tracks from 1973-2015, and (B) sample
tornado tracks for 43 simulation years in Oklahoma City, Oklahoma. .................... 59
Figure 2.19: Comparison between the observed and simulated tornado for the US
by EF scale.................................................................................................................. 61
Figure 2.20: Spatial distribution of observed and simulated tornado by intensity:
(A, B) EF0 and EF1; (C, D) EF2 and EF3; (E, F) EF4 and EF5. ................................. 62
Figure 2.21: Comparison between the observed and simulated tornado spawn
probabilities by month............................................................................................... 63
Figure 2.22: Spatial distribution of simulated and observed tornado by spawn
month: Jan. to Mar. (A, B); Apr. to Jun. (C, D); Jul. to Sep. (E, F); Oct. to Dec
(G, H). ......................................................................................................................... 64
Figure 2.23: Comparison between the observed and simulated tornado spawn
probabilities by hour................................................................................................. 65
Figure 2.24: Spatial distributions of the observed and simulated tornadoes by spawn
hour (CST): 0-5 (A, B); 6-11 (C, D); 12-17 (E, F); 18-23 (G, H) ......................... 66
Figure 2.25: Selected City......................................................................................... 67
Figure 2.26: Probability density functions of tornado spawn hour and spawn month
for (A, B) Indianapolis, Indiana; (C, D) Birmingham, Alabama; (E, F)
Oklahoma City, Oklahoma; (G, H) Des Moines, Iowa............................................ 68
Figure 2.27: Annual spawn frequencies by EF scale for (A) Indianapolis, Indiana (B) Birmingham, Alabama, (C) Oklahoma City, Oklahoma, and (D), Des Moines, Iowa

Figure 2.28: Domain sizes for (A) Moore, Oklahoma, and (B) 50-year tornado wind hazard curves.

Figure 3.1: Illustration of tornado striking probability

Figure 3.2: (A) Birmingham, (B) Oklahoma City, (C) San Antonio, (D) Bozeman

Figure 3.3: Target location (A) and normalized annual probability of exceedance for wind greater than 111 mph (B).

Figure 3.4: location of circle centers and the extent of 15 mi sample circles.

Figure 3.5: Clemson Elementary School, Area ≈ 0.01mi² (A); Greenville Haywood Mall, Area ≈ 0.027mi² (B); Eastman chemical plant in Kingsport, Area ≈ 0.64mi² (C).

Figure 3.6: Flow Chart for determine tornado hazard.

Figure 3.7: Annual probability of experiencing an EF0-EF5 wind speed of point target in the continental United States.

Figure 3.8: Annual probability of experiencing an EF0-EF5 wind speed of small target in the continental United States.

Figure 3.9: Annual probability of experiencing an EF0-EF5 wind speed of medium target in the continental United States.

Figure 3.10: Annual probability of experiencing an EF0-EF5 wind speed of large target in the continental United States.

Figure 3.11: Tornado wind speed hazard for different cities. Oklahoma City, OK(A); Birmingham, AL(B); San Antonio, TX(C); Bozeman, MT(D).

Figure 3.12: Combined tornadoes and non-tornadic extreme wind speed distribution in Oklahoma City (A, B) and Birmingham (C, D).

Figure 3.13: Combined tornadoes and non-tornadic extreme wind speed distribution in San Antonio (A, B) and Bozeman (C, D).

Figure 4.1: Flowchart of tornado track selection and damage estimation procedures.
Figure 4.2: Location of study domain. ................................................................. 114
Figure 4.3: Digitized study domain. ................................................................. 115
Figure 4.4: Distribution of tornado intensity in study domain. ...................... 116
Figure 4.5: Tornado occurrence map across 10,000 years. Tornado tracks of EF0-
EF1 (A); Tornado tracks of EF2-EF3 (B); Tornado tracks of EF4-EF5 .............. 117
Figure 4.6: Modeled tornado impacted area. ................................................. 119
Figure 4.7: Determination of na, nv and nva for a given boundary (V, A). ............. 121
Figure 4.8: MRI for tornado peak wind speed. .............................................. 122
Figure 4.9: MRI for tornado damage area. ................................................. 122
Figure 4.10: Joint MRI for tornado peak wind speed and damage area. ............. 123
Figure 4.11: Image identification and segmentation using MATLAB. .............. 125
Figure 4.12: Identified building boundaries using image identification method. .... 126
Figure 4.13: Buildings in Moore, Oklahoma. ............................................. 127
Figure 4.14: Example of meshed building site. ............................................. 130
Figure 4.15: Example of simulate tornado footprint and damages. ................. 131
Figure 4.16: Tornado caused building damages grouped by different Damage State
(DS). ................................................................................................................. 133
Figure 4.17: Average numbers of damaged buildings caused by three scenarios .... 137
Figure 4.18: Average numbers of damaged buildings caused by three scenarios .... 138
Figure 4.19: Average numbers of damaged buildings in different Damage State
(DS). ................................................................................................................. 139
CHAPTER 1

INTRODUCTION

1.1 Tornado Basics and Background

According to the Glossary of Meteorology (AMS 2000), Tornado is defined as “a violently rotating column of air, pendant from a cumuliform cloud or underneath a cumuliform cloud, and often (but not always) visible as a funnel cloud”.

1.1.1 Tornado spawn location

Tornado can occur in many parts of the world. However, the vast majority of tornadoes are reported in the Central and Southern parts of the United States. “Tornado Alley” (Figure 1.1) is a nickname of the most tornado-prone region in the United States and was first appeared in the title of a research project by the United States (US) Air Force meteorologists Fawbush and Miller (1948). They introduced the name of “Tornado Alley” in their study of severe weathers in an area covering Lubbock, Texas, to Colorado and Nebraska. The phrase is then largely used in the media although many tornado climatologists showed that spatial tornado distributions can vary dramatically depending on the selection criteria. Kelly et al. (1978) studied tornado geographical distribution of path length and of tornado intensity ($F$ scale) and their research shows that the high frequency region for violent tornadoes is not within the tornado alley and the region with long-track tornadoes is located in Mississippi and Louisiana ((Farney & Dixon, 2014)).
This finding is also confirmed by the study of Carbin et al. (2013). Instead of analyzing tornado event, some researchers analyzed the spatial and temporal distribution of the mean number of tornado days per year (Brooks, Doswell, & Kay, 2003; Concannon, Brooks, & Doswell III, 2000). Their result shown that the tornado with an intensity rating of F2 or higher are often observed in Great Plains and Southern United States (Figure 1.2).

![Image of tornado alley]

Figure 1.1: Traditional tornado alley, adopted from NOAA (2016)
1.1.2 Tornado spawn time and season

For the United States as a whole, tornado are most likely occur in months from April to June (Figure 1.3). However, different region may experience tornado “season” at different times of the year. For example, in the region of Gulf coast, tornado are most likely observed during the spring. For the southern Plains, tornado active season is during May into early June. The peak tornado season in the northern plains and Midwest is in June or July (NOAA, 2016; Kelly et al., 1978). But tornadoes can happen any time of the year that favorable condition occurred (Grazulis, 1993).

Figure 1.2: Return frequency of F2+ tornadoes, 1961-2010 (adopted from Carbin et al. 2013)
Because of the diurnal temperature change, tornado occurrence is also highly related to the time of the day (Figure 1.4). Kelly et al. (1978) reveals that the peak frequency occurs at that time of the day when thermal instability is usually the greatest. In other words, tornadoes often spawn during the late afternoon and rarely occur prior to the sunrise.

Figure 1.3: Tornado averages by month 1991-2010 (adopted from NOAA, 2016)
1.1.3 **Tornado path direction**

Common perception of tornado movement direction is toward the northeast. However, distinct seasonal and geographical variation were also revealed by Suckling & Ashley, (2006). According to their study, westerly or northwesterly tornadoes are more likely to happen at late spring and summer in central and northern regions of the United States.
1.1.4 Tornado magnitude

Tornado magnitude or intensity is related to its wind speeds (Table 1.1) and the most common way to rate a tornado is to evaluate the damage it caused. The F-scale tornado rating system was first introduced in 1971 by Fujita and the rating has been adopted by the National Weather Service until 2007. Since then, an improved rating method, Enhanced Fujita (EF) scale, is used to rate tornado intensities in the United States. Compared to F-scale, EF scale has a standardized damage indicator table which considering more types of structures and vegetation, construction quality, and degrees of damage. The EF scale contains six categories from zero to five and each represents a different level of damage. The weakest tornado category, EF0 tornado, means minor or no building damage. The strongest tornado category, EF5 tornado, means total destruction of buildings. In the United States, approximately 85% of tornadoes are EF0 and EF1 tornadoes and the relative frequency decreasing quickly with increasing magnitude, less than 1% are violent tornadoes (EF4 and EF5).

Table 1.1: Tornado intensity and the corresponding wind speed.

<table>
<thead>
<tr>
<th>Magnitude</th>
<th>F-scale wind speed (Mph.)</th>
<th>EF scale wind speed (Mph.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>45-78</td>
<td>65-85</td>
</tr>
<tr>
<td>1</td>
<td>79-117</td>
<td>86-109</td>
</tr>
<tr>
<td>2</td>
<td>118-161</td>
<td>110-137</td>
</tr>
<tr>
<td>3</td>
<td>162-209</td>
<td>138-167</td>
</tr>
<tr>
<td>4</td>
<td>210-261</td>
<td>168-199</td>
</tr>
<tr>
<td>5</td>
<td>262-317</td>
<td>200-234</td>
</tr>
</tbody>
</table>
1.1.5 **Tornado path length and path width**

The relationship between the tornado intensity and the size of the tornado (i.e. path width and length) plays an important role in tornado risk assessment. Long-track tornadoes affect a much larger area and usually cause more fatalities or injuries than short-track tornadoes (Simmons et al., 2011). On average, tornado lengths and widths increase with increasing intensity and this behavior have been modeled using Weibull distribution by Brooks, (2004). Coleman and Dixon, (2014) analyzed tornado path length from the SPC database and they showed that the region with longer tornado path is located in the southeastern part of the US.

1.1.6 **Tornado caused losses and fatalities**

During the period from 1950 to 2015, 60,114 known tornadoes were recorded in the US and these tornadoes caused 933,856 injuries and 5,823 fatalities. On average, 14,149 injuries and 88 deaths are reported every year. Even though there have been a great enhancement of tornado spotting, warning system, and public awareness in recent years, tornadoes still lead to significant damage, injuries and economic losses. During 2001 – 2012, tornadoes were responsible for 1185 deaths (about 16%), ranked the second deadliest weather-related fatalities according to the National Weather Service (NWS). Tornadoes also responsible for almost $25 billion economic losses over the period from 1996 to 2012 (Simmons and Sutter, 2011). After analyzing the tornado report from 1880 to 2005, Ashley (2007) reveals that most killer tornadoes occurred in southeastern US which is not part of
the “tornado alley”. Ashley et al. (2008) examined nocturnal tornadoes during 1880 to 2007, finding that tornadoes that occurred between sunset and sunrise caused more fatalities because most residents were at home or in bed and were unable to take shelter in time because of not hearing the warning of approaching tornadoes at night. Through regression analysis of tornado casualties by different sources, Simmons and Sutter (2005; 2011) found that off-season or winter tornadoes produced more injuries and deaths than tornadoes spawned during the tornado season or summer period, and the highest tornado casualty rate during winter is in southeast region of the US. Two potential explanations of this effect: 1) tornadoes that occurred in the winter often produced heavy snow, ice and widespread severe weather which increase the lethality considerably (Galway & Pearson, 1981), 2) a lulling effect of residents lead them to consider the tornadoes are more likely to occur in the traditional tornado season and tend to ignore or fail to recognize the tornado risks in winter (Simmons and Sutter, 2011).

1.1.7 **Tornado database**

There are two historical archives of tornadoes and both of these databases contain occurrence data, damage classification, and starting and ending location. The Grazulis (1993) database containing over 10,000 tornadoes for the period of 1921-1995. The NOAA database (http://www.spc.noaa.gov/wcm/) contains over 60,000 tornadoes for the period of 1953-present. The major differences between these two databases are the recording time period and the reported tornado intensities, Grazulis database only includes F2 and above tornadoes for a shorter period compared with the NOAA database which contains more
information. It should be noted that the Fujita-scale (F-scale) was first proposed by Dr. Theodore Fujita in 1971 and was not incorporated into both databases until after 1973. Instead of field measurement of tornado intensity based on degree of damage, recorded tornado intensities before 1973 were assigned purely according to the newspaper reports or photographs.

1.1.8 Methodology

The development of techniques for assessing tornado hazard will provide a better understanding of tornado risk and will also give the policy makers or emergency managers more information to make informed decisions. Monte Carlo (MC) simulation approach is a computational method that depends on repeated random sampling from a sequence of probability distributions to obtain numerical result, and this modelling technique have been widely used in risk assessment of relatively rare but high damage consequence natural hazard events, such as flood, hail, hurricane and tornado. Meyer et al. (2002) employed MC modelling approach to study the occurrence distribution of significant tornado (EF2 and higher rate) in the continental US. Daneshvaran and Morden (2007) evaluated the spatial frequency of occurrence of tornadoes in the US and also estimated the losses of tornado and hail outbreaks. Banik et al. (2012) used a stochastic model for assessing the probability of exceedance for the maximum tornado wind speed in southern Ontario, Canada.

Kernel Density Estimation (KDE) method is a non-parametric technique to estimate the probability distribution of a random variable. The probability distribution function can
further be used to determine the probability density at any given location in the study domain rather than only at the exact places where past events occurred. Researchers have employed this method in analysis the spatial pattern for different tornado parameters. For example, Brooks et al. (2003) studied the average annual tornado days in the US. Coleman et al. (2014) analyzed spatial variation of the annual path length of significant tornado in the US. Tan and Hong (2010) evaluated spatial variation of tornado touchdown locations with different intensities in southern Ontario. The value of the bandwidth, or search radius which determines the kernel width at the data point has a strong influence on the estimation result (Dixon et al. 2014). A large bandwidth can cause “oversmoothed” estimation which obscures much of the data variation details and makes estimated densities similar everywhere. In comparison, if a small bandwidth is applied, the estimated density function would contain too many spurious data artifacts and make it very difficult to interpret. Instead of using a plug-in method or any ‘rule of thumb’ method in determining the kernel bandwidth (Banik et al. 2007; Coleman et al. 2014; Concannon et al. 2000; Dixon et al. 2014; Hossain et al. 1999; Widen et al. 2013), Botev et al. (2010) proposed a new estimator which has superior computation efficiency and better performance in the estimation of multimodal density function.

1.2 Objectives

The primary objectives of this study were to:

(1) Development of a stochastic model to simulate tornado tracks in United States.
The stochastic model was developed based on the historical database and includes three sub-models, namely, the genesis, track and wind field models. The genesis model was used to simulate the tornado annual occurrence rate and spawn locations. The track model was utilized to simulate the tornado parameters (such as intensity, heading direction, width etc.) according to its spawn location. The wind field model was used to compute the tornado wind speed along the tornado track. To validate the accuracy of the tornado model, tornado spawn rate and parameters of simulated database was compared to the historical database.

(2) Development of hazard maps in United States for EF0 to EF5 tornado with considering the structure size effect.

A methodology was defined to evaluate the tornado risk for circular target. The influence of reference domain size on estimates local hazard was investigated and the optimized reference domain size was selected. A series of tornado hazard maps for each magnitude considering three building dimensions, namely, small (0.0096 mi²), medium (0.038 mi²) and large (0.62m²).

(3) Quantification of tornado induced property damage on target study region.

Using the simulated tornado database, a new evaluation framework was proposed to accurately simulate the building damages which induced by three selected tornado scenarios. The three scenarios are selected according to the peak wind speed, damage area, and both intensity measures.
1.3 References


CHAPTER 2
TORNADO SIMULATION MODELS

2.1 Abstract

This paper presents the development of a stochastic tornado simulation model for the United States. The continental of the US is subjected to more than 1,000 tornadoes each year, causing significant financial losses and social disruption. Compared to hurricanes, the damage region of a tornado is relatively small and the probability of occurrence at a given location is extremely low. Therefore, it is not feasible to use solely the observed data or tracks to quantify the tornado risk for a given structure or a city that has not been affected by historical tornadoes. In this paper, a methodology for performing stochastic simulation of tornado tracks for the US is presented. The stochastic simulation framework consists of a genesis model, which utilizes the kernel density estimation to simulate the spawn locations of tornadoes. Statistical models for tornado parameters such as track length, path width and intensity, were calibrated using the tornado database maintained by the US National Oceanic and Atmospheric Administration (NOAA) Storm Prediction Center (SPC). The developed statistical models were used to simulate 1,000,000 years of tornado tracks. The simulated tornado parameters include the tornado occurrence rate, intensity (EF scale), location, touchdown time, path length and path width. All these parameters are geographic dependent, meaning the parameters vary depending on the tornado spawn locations. The simulated spawn rates and other key parameters for the continental of the US are compared to the observations. Good agreements are observed between simulations
and observations. To illustrate a potential use of the simulated tornado track database, a probabilistic tornado hazard analysis was performed for Moore, Oklahoma. The 50-year tornado hazard curves for three domain sizes are developed to assess the influence of the domain size on tornado risk.

2.2 Introduction

On average, the continental United States is subjected to more than one thousand tornadoes every year, causing significant financial losses and social disruption. The Storm Prediction Center (SPC), a division of the National Oceanic and Atmospheric Administration (NOAA), maintains a database of tornado events recorded since 1950. The annually observed number of tornadoes, or annual occurrence rate, appears to be increasing (Figure 2.1). This could be due in part to the improvement of technology, such as Doppler radar, used for tracking tornadoes and public awareness in reporting tornado incidents. Compared to a hurricane, the influence area of a tornado is relatively small. Even with over 60,000 of known historical tornado events in the SPC database, many places in tornado prone region have not been hit by a tornado. Therefore, it is not feasible to directly determine the risk due to tornadoes for a location or small region using solely the observed tornado events. In this paper, a stochastic simulation program for generating synthetic tornado tracks based on the statistics of historical data was developed.
Quantifications of tornado hazard and its impact on the built environment are subjects of study by many over the years. Prior tornado climatology research has relied mainly on the spatial and temporal variation of tornado spawn days over a fixed period to quantify tornado risk (e.g. Brooks et al., 2003; Farney and Dixon, 2014). Standohar-Alfano and van de Lindt (2015) divided the continental US into various grid sizes and simulated the annual tornado occurrence probability using the minimum assumption method proposed by Schaefer et al. (1986) in which the tornado occurrence probability is estimated using the sum of the tornado areas divided by the total observation years and the area of the grid of interest. Sigal et al. (2000) simulated multiple realizations of 100,000 years of tornado events for the continental of the US using Latin hypercube method. The simulated

Figure 2.1: Annual number of tornadoes recorded from 1950 to 2015 and 10-year moving average.
results were then used to estimate average annual loss (AAL) for different regions. They concluded that 100,000 years of simulation are not adequate to obtain convergence for AAL. This is likely attributed to very small influence area of tornado.

Boruff et al. (2003) found that while the number of reported tornado events were almost doubled from 1950 to 2000, there has been a steady reduction in tornado induced fatalities and injuries in recent years. This is likely attributed to the advancement made in forecasts and warning times of tornado outbreaks. While the overall fatality rate has reduced, the analyses by Ashley (2007) and Ashley et al. (2008) confirmed the common perception that nocturnal tornadoes caused higher fatalities than tornadoes spawned during the daytime.

Thom (1963) analyzed the distributions of tornado path width and length using tornado data for Iowa and Kansas. He found that more than 90% of the Iowa tornadoes had easterly paths. A more recent study by Suckling and Ashley (2006) examined more than 6000 tornado tracks from 1980 to 2002. They found that while tornadoes generally travel in paths from the southwest toward the northeast direction, in central and northern region of the US, a more westerly tornado paths preponderates during late spring and summer. These studies showed that the spatial and temporal characteristics of tornado paths should be considered. Tan and Hong (2010) developed tornado hazard maps for Southern Ontario in Canada and they also showed that the spatial inhomogeneity of tornado occurrence is an important factor that must be considered when developing tornado hazard maps. In order to simulate the temporal and spatial dependent of tornado tracks, a stochastic simulation
program for generating synthetic tornado tracks based on the statistics of historical data was developed in this paper.

The Monte Carlo Simulation (MCS) technique was employed in this study to develop the tornado simulation program. MCS is a computational method that utilizes repeated sampling of random numbers from a sequence of probability distributions to obtain the behaviors or responses of a relatively complex system or phenomena with random outcomes. The MCS technique has been widely used to assess the risk of natural hazards with relatively rare occurrences. Meyer (2002) employed the MCS approach to study significant tornado occurrence distribution in the continental of the United States. Strader et al. (2016) developed a MCS model for simulating tornado events applied to a user-defined domain to estimate tornado impacts on the built environment. Daneshvaran and Morden (2007) evaluated the spatial frequency of occurrence of tornadoes in the United States and estimated the losses of tornado and hail outbreaks. Banik et al. (2008) used a stochastic model for assessing the exceedance probability of maximum tornado wind speed in Southern Ontario, Canada.

One of the key contributions of the tornado simulation methodology developed in this study is the use of kernel density estimation (KDE) and MCS methods to generate geographic dependent tornado parameters, which include the EF-scale, path length, maximum path width, path direction, spawn month, date, and hours. Many previous studies did not consider the tornado spawn month or time (Daneshvaran and Morden, 2007; Banik et al., 2012; Standohar-Alfano and van de Lindt, 2016), even though the spawn timing of tornadoes has been shown to play an important role in risk assessment. According to the
study by Simmons and Sutter (2010), the fatalities were 15% higher for tornadoes occurred during offshore season compared to tornado season from March to June. In addition, it has been shown that nocturnal tornadoes have higher fatality rate than diurnal tornadoes (Ashley, 2007; Ashley et al., 2008). While advancement in technology and early warning system has greatly reduced the overall number of casualties due to tornadoes, the fatality rates for nocturnal tornadoes remained largely unchanged over the years. Ashley et al. (2008) found that nocturnal tornadoes occurring during midnight to sunrise of local time are 2.5 times more likely to kill than those tornadoes occurring during the day time. Therefore, it is very important to have a model that can explicitly simulate geographic dependent tornado parameters such as EF-scale, path length, path width, spawn month, and spawn time in a day, in particular, when the model is intended for use in estimating occupant risk or casualty.

2.3 Method of analysis

2.3.1 Data sources description and processing

There are two tornado databases that are widely used in tornado related research: (1) the Grazulis database contains over 10,000 tornadoes for the period of 1921–1995; (2) the NOAA database with over 60,000 tornadoes for the period of 1950 to present. Both of these databases contain detailed tornado track information such as, spawn location (latitude and longitude), starting time, width, length, and damage classification. However, the Grazulis database only includes F2 and higher intensity tornadoes prior to 1995. It should
be noted that the well-known Fujita scale tornado intensity classification system was proposed by Dr. Theodore Fujita in 1971 and it was not incorporated into the tornado rating until 1973. The enhanced Fujita (EF) scale was later introduced in 2007. Instead of determining the tornado intensity based on field measurement using the “degree of damage” scale, recorded tornado intensities before 1973 were assigned purely according to the newspaper reports or photographs of the affected regions. Note that both F scale and EF scale are damage based rating system and the numerical categories of both scales are intended to be consistent in terms of the impact or damage to structures. Based on much work from post-tornado field investigations and observations from Doppler radar, the wind speeds associated with the original Fujita scale were deemed too high. This led to the development of the EF scale and the re-assignment of the wind speeds. Since the data set from SPC contains both F and EF scales, a direct mapping of F scale ratings into EF scale is used in this study (e.g., F0 is treated the same as EF0).

Figure 2.1 shows the tornado annual spawn frequency for the continental of the US and the 10-year moving average from 1950 to 2015. The annual spawn frequency in the 1990s increased by 60% compared to the 1950s and increased by 30% compared to 1970s. The observed increased spawn rate in recent decades is likely due to the implementation of Doppler radar network in the early 1990s. In other words, the annual spawn frequency records prior to 1990 may be underestimated.
2.3.2 **Stochastic Track Simulation Model**

The main simulation model includes two sub-models, namely, the genesis model and track model. The genesis model is used to simulate the tornado annual occurrence rate and spawn locations. The track model is utilized to simulate the tornado track parameters (such as intensity, heading direction, width etc.) according to its spawn location. The parameters of each simulated tornado include the intensity (EF scale), touchdown location in terms of the latitude and longitude, touchdown date and time, path length, path width and heading direction.

The parameters of the simulated tornadoes in this study are geographic dependent. In other words, the tornado parameters (e.g. EF scale) are sampled from probability distributions that vary based on geographic location. For example, the likelihood of a major tornado (EF4 or EF5) spawns in Kansas, a tornado prone area, is expected to be significantly higher than that in a location along the eastern coast of the United States. To achieve a geographic dependent simulation, the Kernel Density Estimation method (KDE) is applied in both the genesis model and the track model. The KDE method is one of the most commonly used spatial analytical techniques, which is often used to quantify the spatial variation of the probability density of a random variable. A bivariate normal distribution is utilized to as the kernel density estimator. The probability density function (PDF) of the bivariate normal distribution is:

\[
f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp \left( -\frac{1}{2(1-\rho^2)} \left[ \frac{(x-\mu_x)^2}{\sigma_x^2} + \frac{(y-\mu_y)^2}{\sigma_y^2} - \frac{2\rho(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y} \right] \right)
\]  

2.1
where, $\mu_x$ and $\mu_y$ are the longitude and latitude of the observed spawn location of each tornado. $\sigma_x$ and $\sigma_y$ are the bandwidths for longitudinal and latitudinal directions, respectively, and $\rho$ is the correlation coefficient. The bandwidths represent the likely deviations or drifts of spawn locations of future tornadoes from the known observed locations. In this study, it is assumed that the future spawn locations of tornadoes are equally likely to drift away from the past observed locations in the longitudinal and latitudinal directions, and there is no correlation between the two directions. In other words, $\sigma_x$ is equal to $\sigma_y$ and $\rho$ is taken as zero. As a result, there is only one free parameter to be estimated in Equation 2.1, which is the bandwidth, $\sigma$ (i.e. $\sigma_x = \sigma_y = \sigma$).

The problem at hand is to select an optimal bandwidth for each tornado parameter. An overly large bandwidth may result in over-smoothed estimation, suppressing the actual underlying structure of the probability density distribution. In contrast, if a small bandwidth is applied, the estimated density function may contain spurious statistical artifacts and sharp changes in probability density values between close proximity locations. The bandwidth selection technique via diffusion proposed by Botev and Grotowski (2010) is applied to determine the bandwidth in this study. The selection via diffusion algorithm evaluates the best-fit bandwidth according to the spatial distribution and size of the sample space. Compared to the commonly used mean integrated squared error method, the bandwidth selection via diffusion approach is computational efficiency and it better suited for estimating multimodal density function.
2.3.2.1 **Genesis Model**

According to the previous research by Standohar-Alfano and van de Lindt (2015), negative binomial distribution was determined to be a distribution suitable for describing the tornado annual frequency. The probability density function of the negative binomial distribution is given by:

\[
Pr(X = N_{Tor}) = \binom{N_{Tor} + r - 1}{N_{Tor}} p^{N_{Tor}} (1 - p)^r
\]

where \(N_{Tor}\) is the number of tornados occurred in a year. \(p=0.0253\) and \(r=31.605\) are the distribution parameters fitted using the observed annual spawn rates of the SPC database. The \(p\) and \(r\) parameters were estimated via the maximum likelihood method using the SPC data from 1990 to 2015, and the simulated distribution are shown in Figure 2.2. As previously discussed, the spawn rates prior to 1990 are excluded because it is believed that the dataset may be underestimated due to poor observation coverage.

![Cumulative Density Function (CDF) of tornado annually spawn frequency.](image)

Figure 2.2: Cumulative Density Function (CDF) of tornado annually spawn frequency.
The quantile-quantile (Q-Q) plot is utilized to judge the quality of fit of the observed and modeled probability distributions by plotting their quantiles against each other. If the observed and modeled probabilities have identical distributions, the points in Q-Q plot will approximately lie on a straight diagonal (45-degree) line. In Figure 2.3, the modeled tornado annual occurrence rates are plotted on the x-axis, and the corresponding quantile values from the actual observations are plotted on the y-axis. The red dots in Figure 2.3 represent the 10th, 20th, 30th to 90th percentiles of the two probability distributions. According to the observation, most of the points are close to the 45-degree line which means the fitted negative binomial distribution can be used to model the annual spawn rate of tornadoes in the US. It should be noted that the fitted probability distribution model deviates slightly from the empirical dataset in region of high annual occurrence rates (>1500 tornadoes/year) or beyond the 90th percentile.

Figure 2.3: Quantile-to-Quantile plot of observed versus modeled annual tornado spawn frequencies.
To consider the variation in tornado occurrences due to climatological differences in the US (Farney and Dixon 2014; Kelly et al. 1978), this study modeled the tornado touchdown location as a geographic dependent parameter. Each simulated tornado spawn location is randomly generated using a bivariate normal random number generator with an optimal bandwidth (Equation 2.1) (e.g. see Figure 2.4). The random number generator returns a random location chosen from the bivariate normal distribution with input means ($\mu_x$ and $\mu_y$), and variance ($\sigma$), where the means control the center location of the distribution and $\sigma$ controls the dispersion of the distribution (bandwidth). The means (latitude and longitude) are sampled from the known spawn locations of historical events and variance ($\sigma$) is obtained from the previously discussed KDE by diffusion method.

Figure 2.4: Modeled tornado location by using multivariate normal random number generator.
The tornado genesis model simulation procedures are as follows:

1) Randomly sample a tornado annual spawn rate \( N_{Tor} \) from the negative binominal distribution \( p=0.0253, r=31.605 \);

2) Randomly select a tornado year (1950-2015) and use all the observed tornadoes in that particular year to generate the KDE of the spawn locations with an optimal bandwidth;

3) Randomly select \( N_{Tor} \) spawn locations with replacement using all the observed tornadoes of the selected year in step 2;

4) Use the KDE method to vary the spawn locations determined in step 3 (i.e. use a bivariate normal distribution with the center \( \mu_x \) and \( \mu_y \) equal to the initial spawn locations determined in step 3 and the variance \( \sigma \) equal to the optimal KDE bandwidth determined in step 2 to randomize the final spawn locations).

During the genesis process, if the randomized spawn location of a tornado is outside of the US land boundary (i.e. in ocean), step 4 is repeated until that particular tornado is inside the US land boundary. To preserve the local climatological patterns, instead of aggregating all historical spawn locations to generate one KDE map, a KDE model for spawn location is produced for each simulation year. Figure 2.5 shows an example simulation year with spawn locations of tornadoes derived based on the tornadoes of year 1992 as the seeds. The probability density contours of both the modeled and observed tornado spawn locations are shown in Figure 2.5A and Figure 2.5B. As can be seen, the modeled tornadoes follow the spatial pattern of the observed tornado distribution very well.
Both the observed and modeled tornado spawn KDE contours show high probabilities of occurrence at the northwest region of Kansas and near Louisiana.

Figure 2.5: Probability density contours of tornado spawn locations for (A) modeled, and (B) observed tornadoes using the tornadoes spawned in year 1992 as the seeds.
2.3.2.2  Tracking model

For each tornado spawned using the genesis model, six additional tornado parameters, namely (1) EF scale ($B$), (2) path direction ($\theta$, measured clockwise from the true North), (3) spawn month ($M$), (4) spawn time ($H$), (5) path length ($L$), and (6) width ($W$), are simulated using the track model. To ensure the simulated tornado parameters follow the geographic patterns of historical tornado records, the data for each tornado parameter is divided into subgroups and each subgroup is analysed separately using the same KDE approach employed for the spawn location model. Table 2.1 shows the grouping of the four tornado parameters.

For a specific tornado parameter, for instance the EF scale, the historical tornadoes are categorized into different groups ($B_i$) and each group contains tornadoes with the same characteristic (e.g. EF scale equal to 2). The tornado spawn records from these grouped datasets are used to generate the probability density contour maps using the KDE method. The KDE contours reflect the spatial distribution and concentration of tornadoes with the same characteristic. The developed probability density contour maps are used to determine the point estimate for probability density of tornadoes at a given location with the specified group of parameter of interest. Figure 2.6 shows two examples probability density models (maps) developed using only the EF2 tornadoes (Figure 2.6A) and only those tornadoes spawned during the first half of the month of January (Figure 2.6B).
Table 2.1: Tornado parameter groups.

<table>
<thead>
<tr>
<th>Group $i$</th>
<th>EF ($B_i$)</th>
<th>EF0 ($B_1$)</th>
<th>EF1 ($B_2$)</th>
<th>EF2 ($B_3$)</th>
<th>EF3 ($B_4$)</th>
<th>EF4 ($B_5$)</th>
<th>EF5 ($B_6$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heading Angle ($\theta_i$)</td>
<td>0-22.5 ($\theta_1$)</td>
<td>22.5-45 ($\theta_2$)</td>
<td>45-67.5 ($\theta_3$)</td>
<td>67.5-90 ($\theta_4$)</td>
<td>90-112.5 ($\theta_5$)</td>
<td>112.5-135 ($\theta_6$)</td>
<td></td>
</tr>
<tr>
<td>135-157.5 ($\theta_7$)</td>
<td>157.5-180 ($\theta_8$)</td>
<td>180-202.5 ($\theta_9$)</td>
<td>202.5-225 ($\theta_{10}$)</td>
<td>225-247.5 ($\theta_{11}$)</td>
<td>247.5-270 ($\theta_{12}$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>270-292.5 ($\theta_13$)</td>
<td>292.5-315 ($\theta_{14}$)</td>
<td>315-337.5 ($\theta_{15}$)</td>
<td>337.5-360 ($\theta_{16}$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spawn Hour ($H_i$)</td>
<td>1 ($H_1$)</td>
<td>2 ($H_2$)</td>
<td>3 ($H_3$)</td>
<td>4 ($H_4$)</td>
<td>5 ($H_5$)</td>
<td>6 ($H_6$)</td>
<td></td>
</tr>
<tr>
<td>7 ($H_7$)</td>
<td>8 ($H_8$)</td>
<td>9 ($H_9$)</td>
<td>10 ($H_{10}$)</td>
<td>11 ($H_{11}$)</td>
<td>12 ($H_{12}$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13 ($H_{13}$)</td>
<td>14 ($H_{14}$)</td>
<td>15 ($H_{15}$)</td>
<td>16 ($H_{16}$)</td>
<td>17 ($H_{17}$)</td>
<td>18 ($H_{18}$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19 ($H_{19}$)</td>
<td>20 ($H_{20}$)</td>
<td>21 ($H_{21}$)</td>
<td>22 ($H_{22}$)</td>
<td>23 ($H_{23}$)</td>
<td>24 ($H_{24}$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spawn Date/Month ($M_i$)</td>
<td>Early Jan. ($M_1$)</td>
<td>Late Jan. ($M_2$)</td>
<td>Early Feb. ($M_3$)</td>
<td>Late Feb. ($M_4$)</td>
<td>Early Mar. ($M_5$)</td>
<td>Late Mar. ($M_6$)</td>
<td></td>
</tr>
<tr>
<td>Early Apr. ($M_7$)</td>
<td>Late Apr. ($M_8$)</td>
<td>Early May ($M_9$)</td>
<td>Late May ($M_{10}$)</td>
<td>Early June ($M_{11}$)</td>
<td>Late June ($M_{12}$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Early July ($M_{13}$)</td>
<td>Late July ($M_{14}$)</td>
<td>Early Aug. ($M_{15}$)</td>
<td>Late Aug. ($M_{16}$)</td>
<td>Early Sept. ($M_{17}$)</td>
<td>Late Sept. ($M_{18}$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Early Oct. ($M_{19}$)</td>
<td>Late Oct. ($M_{20}$)</td>
<td>Early Nov. ($M_{21}$)</td>
<td>Late Nov. ($M_{22}$)</td>
<td>Early Dec. ($M_{23}$)</td>
<td>Late Dec. ($M_{24}$)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Once the spawn location of a tornado has been determined using the genesis model, the conditional probability are used to simulate the six tornado parameters. For illustration purpose, consider the determination of EF scale for a tornado $j$ at location $(\text{Lat}, \text{Lon})$. According to the Bayes’ theorem, the probability of a tornado occurs at location $(\text{Lat}, \text{Lon})_j$ and its EF scale is equal to $B$, is:

Figure 2.6: Probability density contours for (A) EF2 tornadoes and (B) tornadoes spawned in early January.
where \( P(B_i) \) is the probability of the EF scale of the tornado is equal to \( B_i \), which can be obtained from Figure 2.7. \( P((\text{Lat, Lon})_j \mid B_i) \) denotes the point estimate for the probability of a tornado spawned at location \((\text{Lat, Lon})_j\) given that the EF scale of the tornado is \( B_i \). To obtain the point estimate probability based on EF scale, a set of probability density maps were developed by grouping the historical tornadoes into groups \( B_1 \) to \( B_6 \) for tornadoes with EF scales equal to 0 to 5, respectively. For instance, the \( P((\text{Lat, Lon})_j \mid B_i = \text{EF2}) \) value for EF2 tornadoes can be obtained from Figure 2.6A.

![Figure 2.7: Probability density function of the contiguous United States tornadoes by EF scale.](image)

Using the simulated tornado spawn location, the intensity of a tornado can then be sampled using a site specific probability density function:
\[ P_{EF,Loc} = \frac{P(EF_i \cap Loc_j)}{\sum_{m=0}^{5} P(EF_m \cap Loc_j)} = \frac{P((Lat, Lon)_j | B_i) \times P(B_i)}{\sum_{m=0}^{5} P((Lat, Lon)_j | B_m) \times P(B_m)} \]

Figure 2.8 shows an example site specific probability density function determined using Equation 2.4 for a location in Oklahoma City. Note that while EF0 tornadoes have the highest occurrence probability for the contiguous of the United States (see Figure 2.7), Figure 2.8 shows that EF1 tornadoes are most likely to spawn in Oklahoma City. A site specific probability density function for EF scale is produced for each tornado based on the spawn location and is used to simulate the EF scale of the tornado.

![Histogram of Probability Density Function](image)

Figure 2.8: Site specific probability density function for a location in Oklahoma City by EF scale.

Similar procedures are applied to simulate the tornado spawn hour, spawn date/month and heading angle. For determining the tornado spawn hour, \( B_i \) represents the tornado spawn hour in a day (1 to 24). Figure 2.9 shows the probability density function for tornado spawn hour for the contiguous US. For the heading angle, the data are grouped into eight equal bins with a 22.5-degree increment. For the spawn date and month, the data are divided into 24 groups with each month split into two segments, first half and second half of the month (Figure 2.10). It should be noted that the first half of each month always
contains 15 days and the second half of the month contains the remaining days of that month. After the tornado spawned month segment has been determined (i.e. early January, late January etc.), the actual spawn day number of the year (1 to 365) is sampled using the occurrence probabilities for those days in the particular month segment using the cumulative distribution function for spawn day of the year as shown in Figure 2.11.

Figure 2.9: Probability density function of the contiguous United States tornadoes by spawn hour.

Figure 2.10: Probability density function of the contiguous United States tornadoes by spawn month.
It has been shown that tornado path length (L) and path width (W) tend to increase with increasing tornado intensity (Brooks, 2004). Therefore, tornado path length and width are simulated according to the EF scale. A more intense tornado tends to have a longer path length and wider path width than that of the weaker ones. Following the study by (Brooks, 2004), the tornado path lengths and maximum widths are modelled using the two-parameter Weibull distribution. The cumulative distribution function of the Weibull distribution is:

\[ F(x) = 1 - \exp\left[-\left(\frac{x}{c}\right)^d\right] \]

where \( c \) and \( d \) are the scale and shape parameters of the distribution. The fitted distribution parameters using maximum likelihood method for path length and path width grouped by EF scale are shown in Table 2.2, and the result are plot in Figure 2.12 and Figure 2.13.

Figure 2.11: Cumulative probability distribution of tornado spawn day of the year (day 1 to 365).
Table 2.2: Modeled Weibull distribution parameters for path length and path width by EF scale.

<table>
<thead>
<tr>
<th></th>
<th>EF0</th>
<th>EF1</th>
<th>EF2</th>
<th>EF3</th>
<th>EF4</th>
<th>EF5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length (km)</td>
<td>c</td>
<td>1.161</td>
<td>4.284</td>
<td>10.28</td>
<td>25.36</td>
<td>44.71</td>
</tr>
<tr>
<td></td>
<td>d</td>
<td>0.6773</td>
<td>0.7288</td>
<td>0.7963</td>
<td>1.037</td>
<td>1.138</td>
</tr>
<tr>
<td>Width (m)</td>
<td>c</td>
<td>41.32</td>
<td>93.56</td>
<td>187.3</td>
<td>414.9</td>
<td>701.8</td>
</tr>
<tr>
<td></td>
<td>d</td>
<td>1.055</td>
<td>0.9431</td>
<td>0.9084</td>
<td>0.9944</td>
<td>1.158</td>
</tr>
</tbody>
</table>

Figure 2.12: Cumulative density function (CDF) of tornado path length.
To ensure that the tornado width and length follow both the statistical distribution and geographic features, the historical tornadoes are firstly grouped according to the EF scale. Then, EF0 to EF4 tornadoes are further divided into four different length (width) sub-group based on the 25th, 50th, and 75th percentiles. Since the data for EF5 tornadoes are limited, the length and width of EF5 tornadoes are split into two sub-groups, divided at the 50th percentile. These grouped dataset are used for generating the probability density contour maps using the KDE method. Table 2.3 and Table 2.4 show the grouping of path length and path width, respectively, based on quantiles and EF scale. For instance, the subgroup $L_2$ in Table 2.3 for EF2 tornado contains all tornadoes with path length that is in

Figure 2.13: Cumulative density function (CDF) of tornado path width.
between 2.2 km and 6.5 km, which correspond to the 25th and 50th percentiles of the EF2 tornado length.

Table 2.3: Tornado length group, $L_i$

<table>
<thead>
<tr>
<th>$L_1$ (km)</th>
<th>$L_2$ (km)</th>
<th>$L_3$ (km)</th>
<th>$L_4$ (km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EF0</td>
<td>PL ≤ 0.2</td>
<td>0.2 &lt; PL ≤ 0.7</td>
<td>0.7 &lt; PL ≤ 1.9</td>
</tr>
<tr>
<td>EF1</td>
<td>PL ≤ 0.8</td>
<td>0.78 &lt; PL ≤ 2.6</td>
<td>2.6 &lt; PL ≤ 6.7</td>
</tr>
<tr>
<td>EF2</td>
<td>PL ≤ 2.2</td>
<td>2.2 &lt; PL ≤ 6.5</td>
<td>6.5 &lt; PL ≤ 15.5</td>
</tr>
<tr>
<td>EF3</td>
<td>PL ≤ 7.6</td>
<td>7.6 &lt; PL ≤ 17.8</td>
<td>17.8 &lt; PL ≤ 34.8</td>
</tr>
<tr>
<td>EF4</td>
<td>PL ≤ 14.9</td>
<td>14.9 &lt; PL ≤ 32.4</td>
<td>32.34 &lt; PL ≤ 59.6</td>
</tr>
<tr>
<td>EF5</td>
<td>PL ≤ 53.4</td>
<td>PL &gt; 53.4</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 2.4: Tornado width group, $W_i$

<table>
<thead>
<tr>
<th>$W_1$ (m)</th>
<th>$W_2$ (m)</th>
<th>$W_3$ (m)</th>
<th>$W_4$ (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EF0</td>
<td>PW ≤ 12.7</td>
<td>12.7 &lt; PW ≤ 29.2</td>
<td>29.2 &lt; PW ≤ 56.3</td>
</tr>
<tr>
<td>EF1</td>
<td>PW ≤ 24.9</td>
<td>24.9 &lt; PW ≤ 63.4</td>
<td>63.4 &lt; PW ≤ 132.3</td>
</tr>
<tr>
<td>EF2</td>
<td>PW ≤ 47.5</td>
<td>47.5 &lt; PW ≤ 125.1</td>
<td>125.1 &lt; PW ≤ 268.3</td>
</tr>
<tr>
<td>EF3</td>
<td>PW ≤ 118.5</td>
<td>118.5 &lt; PW ≤ 286.9</td>
<td>286.9 &lt; PW ≤ 576.2</td>
</tr>
<tr>
<td>EF4</td>
<td>PW ≤ 239.4</td>
<td>239.4 &lt; PW ≤ 511.4</td>
<td>511.4 &lt; PW ≤ 930.4</td>
</tr>
<tr>
<td>EF5</td>
<td>PW ≤ 729.8</td>
<td>PW &gt; 729.8</td>
<td>-</td>
</tr>
</tbody>
</table>

The probability density contour maps conditioned on path length, $P((\text{Lat, Lon})_j | L_i)$, and path width, $P((\text{Lat, Lon})_j | W_i)$, are generated using the same approached used for other parameters such as EF scale and spawn month (see Figure 2.6). The KDE contours maps shown in Figure 2.14 reveal that small-scale tornadoes with lower 25th percentile path length (≤0.33 km (0.21 mi)) and path width (≤14.4 m (15.8 yard)) are often observed in Florida Peninsula, region along the Gulf coast and Central region of the US. Large-scale tornadoes with the path length and path width greater than the third quartile values (PL ≥ 5.5 km (3.4 mi) and PW ≥ 112.9 m (123.5 yard)) are more likely to spawn in Southeast
region of the US (Figure 2.15). Similarly, site specific probability density function for path length and width grouped by quantiles are determined using Equations 2.3 and 2.4. Once the path length and path width subgroups have been determined, the inverse CDF method along with the fitted Weibull distribution parameters shown in Table 2.3 and Table 2.4 are utilized to simulate the tornado path length and width.

![Figure 2.14: Small-scale tornado spawn location density contours for (A) length less than 0.33 km, and (B) width less than 14.4 m.](image-url)
2.3.2.3 Wind field model

2.3.2.3.1 Wind field along the tornado length

The intensity of a tornado along the track usually degrades as the ground friction dissipates the energy of the tornado. After examined 150 tornado tracks, Twisdale and...
Dunn (1981) determined the intensity variation along the track of tornadoes. The fractions of the tornado strength for each of the highest EF scales attained by a tornado are shown in Figure 2.16. In this study, it is assumed that the maximum intensity occurs at the middle of the tornado path and the lower bound of EF0 wind speed, 65 mph, occurs at the fringe of the tornado track.

![Figure 2.16: Tornado intensity variation along the track.](image)

**Figure 2.16:** Tornado intensity variation along the track.

### 2.3.2.3.2 Wind field along the tornado width

The wind speed variation along the tornado width is modeled using the modified Rankine vortex model in which the tangential velocity can be computed as:

\[
V_{\text{tan}}(r) = \frac{r \Gamma_\infty}{\pi (r^2 + r_c^2)}
\]

where \( r \) is the radial coordinate with \( r = 0 \) at the center of the tornado vortex and \( r_c \) is the core radius where the maximum tangential velocity \( (V_{\text{tan},\text{max}}) \) occurs, and \( V_{\text{tan},\text{max}} \) is
assumed to be uniformly distributed between the lower and upper bound wind speeds of the corresponding EF scale. \( \Gamma_\infty = 2\pi r_c V_{\text{tan, max}} \) is the maximum vortex strength. Substitute \( \Gamma_\infty = 2\pi r_c V_{\text{tan, max}} \) into Equation 2.6 yields:

\[
V_{\text{tan}}(r) = \frac{r \Gamma_\infty}{\pi (r^2 + r_c^2)} = \frac{2r (r_c V_{\text{tan, max}})}{r^2 + r_c^2}
\]

where the only unknown in Equation 2.7 is \( r_c \). To solve for \( r_c \), collect the \( r_c \) terms in Equation 2.7 gives the following expression:

\[
r_c^2 - \left( \frac{2r V_{\text{tan, max}}}{V_{\text{tan}}(r)} \right) r_c + r^2 = 0
\]

Note that \( r_c \) is not equal to the maximum path width (\( W \)). According to the SPC database, the maximum path width is a damage based value and this study assumes that the building damage occurs when the tangential wind speed exceeds 65 mph (lower bound wind speed of an EF0 tornado). Assume the lower bound EF0 occurs at the edge of the tornado path width, the core radius \( r_c \) can be determined by setting the tangential wind speed \( V_{\text{tan}}(r) = 105 \text{ km/h} \) (65 mph.) at \( r = W/2 \):

\[
r_c^2 - \left( \frac{W}{105} V_{\text{tan, max}} \right) r_c + \left( \frac{W}{2} \right)^2 = 0
\]

The core radius is computed by substituting the simulated path width (\( W \)) and maximum tangential wind speed into Equation 2.9.
2.4 Stochastic Tornado Track Simulation Program

A computer program for stochastic simulation of tornado tracks has been developed and coded using Matlab computing program language. The organization of the simulation program is shown graphically in a flow chart in Figure 2.17. The modules for the Genesis Model are marked with purple color and the modules for the Track Model are marked in orange color. The simulation steps for the program are as follows:

Figure 2.17: Flowchart for stochastic tornado track simulation program.

...
(1) Input the total number of simulation years \((N_{\text{year}})\).

(2) For each simulation year \((i = 1, 2, 3, \ldots N_{\text{year}})\), determine the number of tornadoes \((N_{\text{tor}})\) in year \(i\) using the negative binomial distribution defined in section 2.3.2.1.

(3) For each modeled tornado \(j (j = 1, 2, 3, \ldots N_{\text{tor}})\), generate spawn location using the genesis model defined in section 2.3.2.1.

(4) Generate the spawn month, spawn date, spawn hour, path direction, EF scale, length and width based on the procedures discussed in section 2.3.2.2.

(5) Record the simulated tornado information and repeat steps 3 and 4 if \(j \leq N_{\text{tor}}\) (i.e. simulate the parameters for each tornado in the \(i\)-th year).

(6) Repeat steps 2 to 5 if \(i \leq N_{\text{year}}\) (i.e. repeat for \(N_{\text{year}}\) simulation years)

### 2.5 Synthetic Tornado Tracks Database and Applications

The developed computer program is utilized to generate 1,000,000 years of synthetic tornadoes. The simulated database contains more than 1 billion simulated tornado tracks. To verify the applicability of the simulated tracks, comparisons are made between the simulated and observed tornadoes for intensity, spawn month and spawn hour for various locations and the verifications are presented in the next sections. In addition, using the catalog of simulated tornado tracks, a probabilistic tornado hazard analysis is performed for Moore, Oklahoma. To study the influence of domain size on tornado risk, tornado hazard curves for three different domain sizes are generated for a location in Moore, Oklahoma.
2.5.1 **Tornado tracks**

As an illustrative example, comparison between the simulated and observed tornado tracks within a radius of 40 mi (64.4 km) from the Oklahoma City for an observation period of 43 years (1973 to 2015) is shown in Figure 2.18. The full paths of the tornadoes are shown in Figure 2.18 along with the EF scale identified by color. The simulated tracks visually agree with the patterns of the historical tornado tracks. The corresponding tornado counts for each EF scale are shown in Table 2.5, which match the historical counts reasonably well. Note that Table 2.5 shows the results for one realization over a 43-year time frame. The results may vary for different realization.

![Figure 2.18: (A) Actual observed tornado tracks from 1973-2015, and (B) sample tornado tracks for 43 simulation years in Oklahoma City, Oklahoma.](image)

Table 2.5: Number of simulated and observed tornado tracks near Oklahoma City

<table>
<thead>
<tr>
<th></th>
<th>EF0</th>
<th>EF1</th>
<th>EF2</th>
<th>EF3</th>
<th>EF4</th>
<th>EF5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Observed</strong> (1973-2015)</td>
<td>120</td>
<td>124</td>
<td>41</td>
<td>20</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td><strong>Simulated</strong> (43 years)</td>
<td>129</td>
<td>115</td>
<td>39</td>
<td>19</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>
A previous study by Sigal et al. (2000) has shown that 100,000 simulation years may not be adequate to achieve stability of the simulated tornado hazard. A convergence study was carried out to determine the stability of the simulated occurrence rate for multiple realizations of 1 million simulated years for a 2-mile circular study domain located in Oklahoma City, Oklahoma. Figure 2.19 shows the convergence plot of the coefficient of variation (CoV) of the spawn rate versus simulation year. It can be seen that about 200,000 simulation years are needed to keep the CoV of spawn rate of all tornadoes (i.e. EF0 and higher) to less than 0.01. For intense tornadoes (EF4 and higher) that are more rare, slightly less than 1 million simulation years are needed to maintain the CoV of spawn rate to less than 0.01.

Figure 2.19: convergence of spawn rate for Oklahoma City versus simulation years.
2.5.2 **Tornado intensity**

Comparison between the simulated and observed tornado probability density functions for the contiguous US by EF scale are shown in Figure 2.20. The breakdown of the simulated tornadoes by EF scale matches the past observations very well, which confirms that the simulation program produces the correct ratios for different EF scale tornadoes.

![Figure 2.20: Comparison between the observed and simulated tornado for the US by EF scale.](image)

Probability density maps and contours are generated for simulated and observed tornadoes to investigate the spatial variation of tornado by intensity (Figure 2.21). The patterns of the probability density contours of the simulated tornadoes for each intensity group match that of the contours from the observed tornadoes. Weak tornadoes (EF0 and EF1) have a wide spread area of occurrences and they cover the midsection and Southeast portion of the US. In addition, except for those occurred in the Tornado Alley, the weak tornadoes are also likely to spawn in Florida peninsula and region around the Gulf coast. The high occurrences of weak tornadoes in the coastal regions are likely caused by
additional tornadoes that spawned during the landfall of tropical cyclones or hurricanes. Strong tornadoes (EF2 and EF3) have high probabilities of occurrence in the Southeast region of the US, which includes portion of the Tornado Alley and most of the Dixie Alley. The peaks of the probability density contours for major tornadoes (EF4 and EF5) are observed in Tornado Alley, Dixie Alley and Midwest. More details about the spatial variation of tornado intensity can be found in Appendix D.

Figure 2.21: Spatial distribution of observed and simulated tornado by intensity: (A, B) EF0 and EF1; (C, D) EF2 and EF3; (E, F) EF4 and EF5.
2.5.3  

**Tornado spawn month**

The seasonal variations of the occurrence probabilities of tornadoes are explicitly considered in this study. The comparison between the simulated and observed tornado monthly spawn rates is shown in Figure 2.22, which shows very good agreements between the simulated and observed spawn probabilities for all twelve months.

![Figure 2.22: Comparison between the observed and simulated tornado spawn probabilities by month.](image)

Due to strong wind shears and atmospheric instability that often occurs in spring and summer, the months with high tornado spawn probabilities are April to July. The geographic and seasonal dependent behaviors of simulated tornadoes are shown in Figure 2.22. The geographic regions with high spawn probabilities change dramatically with the change in season. During the winter season (e.g. see Figure 2.6, early January), tornadoes generally spawn in the Southeast region where as during the summer, tornadoes generally occur in the Mid-west and North Plains. More details about the spatial variation of tornado spawn month can be found in Appendix B.
2.5.4 Tornado spawn hour in a day

It has been shown that tornado occurrence is highly correlated to the time/hour in a day and the density closely follows the diurnal temperature curve (Kelly et al., 1978), with
the peak occurrence appears during the late afternoon, while minimum occurrences just prior to the sunrise (Figure 2.24).

![Figure 2.24: Comparison between the observed and simulated tornado spawn probabilities by hour.](image)

The spatial distribution contours of tornado occurrences grouped by hour in a day look similar to the contours of tornado occurrence time in a year. When the diurnal temperature is low, between 6 am (CST) to 12 am (CST), tornadoes are most likely observed in Florida and along the Gulf Coast (Figure 2.25). The peak spawn locations move to Midwest and further spread out into the North and Northeast regions of the US between 12 am (CST) to 11 pm (CST). Finally, the regions with peak occurrence probability return back to Florida, the Gulf Coast and Midwest areas between 12 pm (CST) to 5 am (CST). More details about the spatial variation of tornado spawn time in a day can be found in Appendix C.
Comparisons for Select Cities

Comparisons are made between the simulated and observed statistics for tornado intensity, spawn month and spawn hour for select cities. Four cities are chosen based on
the degrees of tornado activity in the regions (Figure 2.26: Selected City.). These cities are: (1) Des Moines, Iowa (41.577, -93.617) located in High Plains region; (2) Oklahoma City, Oklahoma (35.457, -97.514) located in tornado alley; (3) Indianapolis, Indiana (39.777, -86.148) located in Midwest region; (4) Birmingham, Alabama (33.536, -86.798) located in Dixie alley. A search radius of 40 km (25 mi) from the city center is used to identify the tornadoes that affected the city of interest. Those tornadoes within the search radius are used to compare the statistics of the simulated and actual observed tornadoes.

Figure 2.26: Selected City.
2.5.5.1  Tornado spawn month/time

The probability mass functions of tornado spawn hour and month for the selected locations are shown in Figure 2.27. The red x marks are the means of the observed probability density values and the blue bars are the simulated probability densities. The 95% confidence intervals are plotted as red lines. It can be seen from Figure 2.27 that the
probability densities of the simulated tornado spawn time and month match the observed data. All cities experience high tornado activities from March to June. Some study have shown that, nocturnal tornado is the main reason that causes high fatality rate in the South-eastern region of the United States (Ashley, 2007; Ashley et al., 2008). Unlike other cities which tornado often observed during the afternoon, Figure 2.27 shows that Indianapolis and Birmingham have relatively high probabilities of observing tornadoes during the night.

2.5.5.2 Tornado Intensity

The annual tornado spawn frequency within a radius of 40 km (25 mi) of the selected cities for different EF scales are plotted in Figure 2.28. The histograms show the simulated tornado frequencies and the x markers are the means annual frequencies of historical events. Also shown in Figure 2.28 are the 95% confidence intervals of the annual spawn rates estimated based on limited historical tornado events. The simulated track yields very good simulation results in those selected cities, with all the mean of simulated annual frequencies within the 95% confidence intervals. The number above the histogram are the mean number of simulated tornado tracks (top) and observed tornado tracks (bottom) for an observation period of 66 years. Note that there were no EF5 tornado reported in Indianapolis and Des Moines over the record period (1950 to 2015); however, this does not mean the occurrence probability of EF5 tornadoes in these cities are zero. Based on the simulation program, the model predicted annual spawn frequencies for EF5 tornadoes for Indianapolis and Des Moines are $3.6\times10^{-3}$ and $5.2\times10^{-3}$, respectively.
Figure 2.28: Annual spawn frequencies by EF scale for (A) Indianapolis, Indiana (B) Birmingham, Alabama, (C) Oklahoma City, Oklahoma, and (D), Des Moines, Iowa.
2.5.5.3 **Hazard curves**

To demonstrate one of the many potential applications of this simulated tornado database, tornado hazard curves in terms of 50-year exceeding probability versus maximum wind speeds are developed for Moore, Oklahoma for three different domain sizes (Figure 2.29). The maximum wind speeds occurred inside the domain are computed using the wind field model presented in section 2.3.2.3. The radii of the three circular domains considered are 0.16 km (about 0.1 mi, small domain), 1.6 km (about 1 mi, medium domain) and 3.2 km (about 2 mi, large domain). All three domains are centred in Moore, Oklahoma and the study is aimed at assessing the effect of domain size on tornado hazard curve.

![Diagram showing domain sizes and hazard curves](image)

**Figure 2.29**: Domain sizes for (A) Moore, Oklahoma, and (B) 50-year tornado wind hazard curves.
Research has shown that tornado occurrence rate can be reasonably simulated by a Poisson process and that the probability that the peak wind speed $v_i$ is larger than a certain wind speed value $V$ induced by tornado during time period $T$ can be described as:

$$P_T(v_i > V) = 1 - \exp \left( -\frac{n}{Y} T \right)$$

where $n$ is the total number of tornadoes producing wind speed of greater than the threshold value $V$ inside the study domain. $Y$ is the total number of simulation years (i.e., 1 million years in this study). $T$ is taken as 50 years for 50-year hazard curve:

$$P_{50}(v_i > V) = 1 - \exp \left( -\frac{n}{Y} \times 50 \right)$$

The 50-year hazard curves in Figure 2.29 show the domain size effect. The results indicate that the 50-year exceedance probability increases with increasing domain size and decreases with increasing peak wind speed. Based on the hazard curves shown in Figure 2.29, there is a 95% probability that the large domain in Moore, Oklahoma will experience at least one EF0 tornado with wind speed exceeding 105 km/h (65 mph) in a 50-year time span, whereas the small domain has about 9.8% chance of experiencing EF0 or stronger tornadoes. The hazard curves for various domain sizes may be used by engineers to design for structures to resist tornado loading.

2.6 Summary and Conclusion

In this study, the NOAA SPC tornado database is utilized to develop a stochastic simulation program. The spawn or touchdown locations are simulated using geographic dependent kernel density estimation (KDE), which specifically accounts for the spatial
distribution of tornadoes properties at different geographic regions (e.g. tendency to spawn strong in Dixie alley and etc.). The simulated track parameters include the tornado occurrence rate, intensity (EF-scale), touchdown location, touchdown time and path direction. All these parameters are geographic dependent, meaning the properties vary depend on the geographic locations. The simulated spawn rates and other parameters for the contiguous US and for four select cities are compared to observations and the modeled results compared well with the observed tornado records. As an illustrative example, the 50-year tornado hazard curves for Moore, Oklahoma with three domain sizes are generated using the simulated tornado database. The results show that domain size has a significant influence on the tornado hazard curve. Therefore, size effect (e.g. single-family versus big box store) may need to be considered in building code for tornado design. The developed tornado database may be used by engineers for performance-based design or risk analysts for catastrophe modeling and loss estimation for tornado hazards.

2.7 References


CHAPTER 3
TORNADO HAZARD ANALYSIS AND EFFECT OF STRUCTURE SIZE

3.1 Abstract

The United States of America experiences more than 1000 tornadoes every year. Different from other large scale natural hazards such as earthquake and hurricane, the impact of a tornado is relatively small. The effect of structure size on tornado risk assessment is very important. Neglecting the structure size may lead to underestimation of tornado strike probability. This study presents the development of size-dependent tornado design maps for the United States. Using a stochastic tornado simulation model and a wind field model, tornado hazard maps for EF0 to EF5 wind speeds are developed for four different target structure sizes, namely point target, small (0.08 mi$^2$), medium (0.03 mi$^2$) and large (0.5 mi$^2$) circular targets. A model to quantify the relationship between tornado strike probability and target size is proposed. Using this relationship, the tornado hazard for a given location and structure size can be interpolated from the four size dependent hazard maps. This quantitative tornado hazard estimation method considering size effect can be used by engineers to determine the design wind speed.
3.2 Introduction

Unlike a hurricane, the footprint of a tornado is relatively small. While tornadic wind can be more violent than hurricane wind, due to the relatively small spatial coverage of tornadic wind, tornado is considered a low probability and high consequence event. To determine the risk of building stock exposed to potential tornado devastation, hazard maps which accurately estimate the tornado striking probability are deemed necessary (Boruff et al., 2003; Meyer et al., 2002; Sigal et al., 2000; Standohar-Alfano et al., 2014; Strader et al., 2016; Tan et al., 2010; Thom, 1963). Unlike hurricane or typhoons, tornados are short lived and localized events. Due to the unique characteristics of tornado risk, the effect of structure size plays an important role on risk calculation and damage assessment (S. Banik et al, 2007; S. Banik et all, 2008; Ramsdell et al., 2007; L. A. Twisdale et al., 1983). Neglecting structure size may result in significant underestimation of tornado strike probability for structures with large area footprint or large-scale infrastructure. Therefore, size effect must be considered when evaluating tornado risk for critical infrastructure with a large spatial coverage area, such as school, hospital, nuclear power plant or petrochemical plant.

Tornado risk assessment has received considerable attention over the last few decades; however, size effect was neglected in many of the past studies (Meyer et al., 2002; Romanic et al., 2016; Standohar-Alfano et al., 2014; Strader et al., 2016; Tan et al., 2010). Size effect for one dimensional (1-D) line structures such as electric power transmission lines has been addressed by (S. Banik et al., 2008; L. A. Twisdale et al., 1983) Buildings and other structures with two dimensional (2-D) footprints cannot be modeled as line
structures. For engineering design and risk evaluation purposes, high resolution tornado hazard maps which cover the whole continental United States are needed. The tornado maps developed in many past studies utilized reference domain with a 1-degree or higher grid resolution (one degree latitude is approximately 69 miles (111 km) apart). These coarse resolution hazard maps may obscure the risk variation details in small region.

The main objectives of this study were: (1) to define and evaluate the tornado risk for 2-D structures, (2) to determine the appropriate grid spacing for high resolution tornado hazard maps, (3) to generate a series of tornado hazard maps for different intensity (EF scale) and structure sizes.

3.3 Tornado track database

To perform tornado hazard analysis, a database of past known tornado events is needed. The US National Oceanic Atmospheric Administration (NOAA) Storm Prediction Center (SPC) has compiled a database of past historical events, with more than 60,000 tornado recorded since 1953. The annually observed number of tornadoes, or annual occurrence rate, appears to be increasing. This could be due in part to the improvement of technology used for tracking tornadoes and the public awareness in reporting tornado incidents. Even with more than 60 years of data with over 60,000 known tornado events, there are many areas in the US that have not been hit by tornadoes or do not have any official record. Hence, it may not be feasible to estimate tornado risk solely based on past observations especially for high resolution risk assessment.
In order to estimate the risk for regions that have not been hit by historical tornadoes, a stochastic tornado track simulation method proposed by Fan et al. (2017) was applied in this study (see Chapter 2). A simulated tornado database with one million years of tornado tracks was generated using the stochastic simulation model. Each simulated track includes tornado parameters, such as intensity (EF scale), spawn location, touchdown time, path length and path width. The tornado track parameters are geographic dependent, meaning the parameters vary based on the tornado spawn locations.

3.4 Tornado hazard for point and area

3.4.1 Tornado Striking probability

Thom (1963) proposed a method to estimate the probability of tornado striking a point, and the equation is expressed as:

\[ P(V \geq v| Tor) = \frac{A_{Sl}}{A_R} \]

where \( P(V \geq v| Tor) \) estimate the probability of tornado striking a point with maximum gust wind speed \( V \) greater than a given value \( v \). For a point target, \( A_{Sl} \) is defined as a tornado covered area which \( V \geq v \); and \( A_R \) is the tornado reference area or region.

For the striking probability of circular area,, the Equation 3.1 mentioned above is still applicable, however \( A_{Sl} \) have to redefine as shown in Figure 3.1. In such case, for a circular area with radius \( r \), \( A_{Sl} \) is the tornado covered area plus the area paint in yellow. The yellow region is a region where tornado strike occurs if the circular target center lies
within. Based on this definition, \( r \) is approaching to zero as the target structure size getting smaller. \( r \) equals to zero when the target is a point and \( A_{si} \) is equal to the area enveloped by tornado track.

![Figure 3.1: Illustration of tornado striking probability](image)

The above discussion is regard to the striking probability for a single tornado track, in order to assess the annual striking probability by using the simulated tornado tracks database, the procedures have been shown as follow. If \( v = 65 \), which is the lower bound of EF0 tornado. The probability of a point not affected by a tornado strike is defined as:

\[
P(V < 65 \text{ mph.} | Tor_i) = \frac{A_R - A_{si}}{A_R}
\]

probability of tornado not striking a point for year \( j \) \((P_{j,NS})\) is:

\[
P_{j,NS} = \prod_{j=1}^{n} P_j(V < 65 \text{ mph.} | Tor_j)
\]

probability of tornado striking a point for year \( j \) \((P_{j,S})\) is:
\[ P_{j,S} = 1 - P_{j,NS} \]

then, annual probability of tornado striking a point with maximum gust wind speed \( V \geq v \)

\[
P(V \geq v | Tor_i)_{Annual} = \frac{\sum_{j=1}^{N_{year}} P_{j,S}}{N_{year}}
\]

where \( N_{year} \) is the total simulation years and \( N_{year} \) equals to one million years in this study.

3.4.2 Reference domain for the uniform hazard

The tornado hazard varies depends on geographic location. For instance, the annual occurrence rate for a location in tornado valley is expected to be higher than a location along the eastern coast of the US. While the tornado hazard may vary in a large geographic region, it is assumed that the tornado risk remain uniform (or approximately uniform) within a small region. The most appropriate reference domain size (i.e. grid spacing for hazard map) should accurately reflect the spatial variation of local tornado hazard. A large reference domain can cause “oversmoothed” effect for hazard map which obscures much of the risk variation details in small region. To determine the optimal domain size, tornado striking probabilities for a point structure are evaluated at four locations with a varying circular reference domain radius range from 65.5 feet to 25.5 miles.

The four locations were selected from different regions of the United States based on the tornado risk level. The first two locations were selected in high risk region, Birmingham located in the north central region and Oklahoma City located in the Southern
Great Plains region (Figure 3.2, (A) and (B)). The other two locations were selected in moderate or low risk region, San Antonio located in the South Central region and Bozeman located in the North-western United States (Figure 3.2, (C) and (D)). At each location, tornado strike probabilities for a point-like target with varying reference domain size were calculated using the method mentioned in section 3.4.1. In high risk region, strike probability of weak (EF0 and EF1) and strong (EF2-EF3) tornadoes are not very sensitive to the change in domain size, mainly because there are sufficient number of simulated tornado events in that region. However strike probability of violent tornadoes (EF4-EF5) may fluctuate if the domain radius is less than 10 miles because the simulated database (1,000,000 years) is not long enough to capture this kind of rare event in such a small region.

Similar patterns were also observed for moderate and low risk regions. If the reference domain radius is less than 10 miles, the estimated annual strike probabilities of EF3 to EF5 tornadoes for San Antonio, Texas (Figure 3.2, C) and Bozeman, Montana (Figure 3.2, D) show variation. In order to maintain a balance between computation cost and accuracy, a 15-mile radius is selected as the reference domain with approximately uniform hazard and the grid points for the tornado hazard maps are spaced 15 miles apart.
3.4.3 **Target size effect**

It has been determined that the tornado hazard is approximately uniform within a 15-mile radius. However, the risk of a structure may vary depending on the size of the structure even within a 15-mile uniform hazard region. To investigate the size effect of...
structure on tornado risk, seven different target domain (structure) sizes ranging from 0.025 mi to 0.8 mi (Table 3.1). The area of the largest target size is about 1024 times larger than the smallest target. All target domains are assumed located near Oklahoma City, and the location is shown Figure 3.3 (A). The Will Rogers World Airport (OKC) which has a footprint of about 0.47 mi² and the Moore High School which has a footprint of about 0.025 mi² (688,596 ft²), are used to establish the domain sizes considered in this study. The size of the OKC airport is between target 5 (0.63 mi²) and target 6 (1.22 mi²), while the size of the Moore high school is in between target 2 and target 3.

The tornado risk for these 6 targets have been investigated by using the method in section 3.4.1. Figure 3.3 (B) clearly shows that tornado hazard for nonzero structure size target can be several orders of magnitude higher than that for a point-like target. For instance, the annual probabilities of exceeding a 65 mph wind (lower bound wind speed of EF0 tornado) are about $2.1 \times 10^{-2}$ and $7.5 \times 10^{-4}$ for a target of 2.43 mi² and a point-like target, respectively. The MRI for observing tornadoes with wind speed exceeding 65 mph are 47 years and 1323 years for a target of 2.43 mi² and a point-like target, respectively. The increase in the probability of exceedance for a structure with a finite size target is highly dependent on the size. Such a relationship can be expressed using a power function:

$$y = ax^b + c$$  \hspace{1cm} (3.6)

where $a$ serves as a scaling factor; $b$ is the exponent; $c$ is the y intersection. The fitted curve is plotted in Figure 3.3 (B). It shows that for a small target with an area of about 0.0024 mi² (67,000 ft²), the strike probability is about 2 times higher than a point-like structure.
For large target with an area of $2.43 \text{ mi}^2 (6.7 \times 10^7 \text{ ft}^2)$, the strike probability is about 30 times higher than a point-like structure.

Table 3.1: Target size information

<table>
<thead>
<tr>
<th>Number</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Radius</td>
<td>0.025</td>
<td>0.07</td>
<td>0.14</td>
<td>0.28</td>
<td>0.40</td>
<td>0.56</td>
<td>0.8</td>
</tr>
<tr>
<td>Area</td>
<td>2.4e-3</td>
<td>1.9e-2</td>
<td>7.6e-2</td>
<td>0.31</td>
<td>0.63</td>
<td>1.22</td>
<td>2.43</td>
</tr>
<tr>
<td>Scale</td>
<td>1</td>
<td>8</td>
<td>32</td>
<td>128</td>
<td>256</td>
<td>512</td>
<td>1024</td>
</tr>
</tbody>
</table>
Figure 3.3: Target location (A) and normalized annual probability of exceedance for wind greater than 111 mph (B).
3.5 Simulation procedures

3.5.1 Setup stations

Figure 3.4: location of circle centers and the extent of 15 mi sample circles.

For this study, we perform the simulation using overlapping 15 mi radius circles centered on a 15 mi spacing grid spanning from 67° W to 124° W, and 25° N to 49° N. The grid points covered the whole continental US and there are 17,646 points in total, the location has shown in Figure 3.4. A more detailed zoom-in view for Oklahoma City is also presented in Figure 3.4. As illustrated in the figure, a 15-mi radius circular area is used to generate the hazard curve for each grid. Note that this approach allows overlapping of region with neighboring grid points.
3.5.2 Determine target size

According to the study of Twisdale (1983) and the discussion in section 3.4.3. The effect of structure size plays an important role in tornado wind speed risk analysis. In order to consider the size effect, three circular targets with different sizes are used in the mapping of tornado hazard. The area of the first (small) target is about 0.0096 mi\(^2\), the area of the second (medium) target is about 0.038 mi\(^2\), and the area of the third (large) target is about 0.62 mi\(^2\). For easy reference, these three targets are termed small target, medium target and large target, respectively (Table 3.2).

<table>
<thead>
<tr>
<th></th>
<th>Small Target</th>
<th>Medium Target</th>
<th>Large Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>Area ((mi^2))</td>
<td>(\approx 0.0096)</td>
<td>(\approx 0.038)</td>
<td>(\approx 0.62)</td>
</tr>
<tr>
<td>Area ((km^2))</td>
<td>(\approx 0.020)</td>
<td>(\approx 0.077)</td>
<td>(\approx 1.3)</td>
</tr>
</tbody>
</table>

Table 3.2: Target size applied in hazard maps
Small target was defined to consider the tornado hazard for school or hospital, and Clemson Elementary School with an area of 0.01\textit{mi}^2 was shown in Figure 3.5 (A) as an illustrative example. Medium target has a similar size to large commercial building or shopping mall, for example, the Greenville Haywood Mall (Figure 3.5 (B)) is in this category which has the area of 0.027 \textit{mi}^2. The size of large target is commonly observed in airport, nuclear power plant and petrochemical facility. The Eastman chemical plant in Kingsport is conducted as an example of large target in Figure 3.5 (C). It has to note that

Figure 3.5: Clemson Elementary School (SC), Area \( \approx 0.01mi^2 \) (A); Greenville Haywood Mall (SC), Area \( \approx 0.027mi^2 \) (B); Eastman chemical plant in Kingsport (TN), Area \( \approx 0.64mi^2 \) (C).
the shape of these targets have idealized as a circular domain, structure with a high length-to-width aspect ratio is not covered in this study, for example, power transmission lines. However, pervious study (S. Banik et al., 2007; S. Banik et al., 2008; Savory et al., 2001) have shown that the effect of orientation and the dimension of transmission line is significant if the length of a line structure greater than 1,000 m. These three types of target are also observed in tornado damage survey report. Study of Matsangouras (2010) shown the significant damage of tornado event in Athens to the local airport and the parked airplanes. Severe damage of Plaza Tower Elementary School in Moore (OK) was caused by Tornado impact in 2013 (Kuligowski et al., 2013).
3.5.3 Analysis procedure

The analysis flow chart is illustrated in Figure 3.6, and the major simulation steps for the analysis are as follows. In order to clearly reflect the size effect in tornado hazard analysis, there are three target size applied in this study. For each target size,
1. Load tornado database. The simulated tornado database include 1,000,000 years of simulation ($N_{year}$) and the simulation procedures have been explained in previous chapter.

2. For study domain $j$ ($j = 1, 2, 3, \ldots N_{domain}$), determine domain envelop and area ($A_{r,j}$).

3. For simulation year $t$, determine total number of tornados that hit domain $j$.

4. For each selected tornado ($i$), calculate tornado impact area $A_{st}$. Determine the non-striking probability of current tornado based on the procedures defined in section 3.4.1.

5. Save the non-striking probability of each impacted tornado based on the equations discussed in section 3.4.1 and repeat step 4 until $i = N_{tor,jt}$.

6. Calculate and record tornado striking probability of year $t$ for domain $j$ according to the equations discussed in section 3.4.1 and repeat steps 3 to 5 if $t \leq N_{year}$.

7. Calculate and record tornado annual striking probability of domain $j$, and repeat steps 2 to 6 if $j \leq N_{domain}$.

3.6 Result and Discussion

3.6.1 Tornado hazard maps

Tornado strike probability for point and area targets are estimated using the simulated tornado database (1,000,000 simulation years) at every grid point in the United States. A hazard map for a given target size is then created for each of the five tornado
intensity levels (EF0 to EF5). The simulated tornado hazard maps for a point target and circular target with different sizes are presented from Figure 3.7 to Figure 3.10. Figure 3.7 present contour maps of tornado striking probability for point-like structure for continental United States.

Figure 3.7(A) to Figure 3.10(A) show the annual probability (P) of experiencing EF0 and higher wind speed \( V \geq 65 \text{mph} \) for different target sizes. The high risk regions (\( P = 10^{-3} \) to \( 10^{-2} \)) of experiencing EF0 and greater tornado are located in the Tornado Alley (extends from northern Texas, Oklahoma, Kansas, into Nebraska) and Dixie Alley (stretches from eastern Texas and Arkansas across Louisiana, Mississippi, Tennessee, Alabama, Georgia, to upstate South Carolina, and western North Carolina). The low risk tornado regions (\( P = 10^{-9} \) to \( 10^{-10} \)) for EF0 and greater tornadoes are observed in the Western US, between the Rocky Mountains and the West Coast of the US.

The annual striking probability is increased by increasing the target size, the order of magnitude of high risk region is up to \( 10^{-2} \) when the large target is applied. Also, attention needs to be given to the West Coast of United States, especially California. Although tornadoes occurred in this region are mostly weak tornadoes (from EF0 to EF2), the annual probability of experience tornado strike could up to \( 10^{-3} \) based on the size of the structure. It should be noted that West Coast often has an irregular tornado season as compared to the tornado occurred in Tornado Alley or Dixie Alley, and California’s tornado season is primarily during January to March which has been discussed in Section 2.5.3.
The annual probability of experiencing EF1 and greater \((V \geq 86\text{ mph})\) tornadoes for target with different size are presented in Figure 3.7 (B) to Figure 3.10 (B). Tornado hazard in this category covers all contiguous states, and the geographic pattern is nearly identical to the maps of EF0 tornado. The low risk \((10^{-10})\) region include most areas of West United States, and high risk region \((10^{-3})\) is located in Tornado Alley and Dixie Alley.

The annual probability of experiencing EF2 and greater \((V \geq 111\text{ mph})\) tornadoes are presented in Figure 3.7 (C) to Figure 3.10 (C) for point target, small target, medium target and large target, respectively. The highest order of magnitude of tornado hazard has reduced to \(10^{-4}\) compared to EF0 and EF1 maps. The spatial distribution of high risk regions tend to follow a similar pattern of pervious maps. However the coverage of low risk regions were extremely increased, because of the number of EF2 to EF5 tornado only accounts 20% of database. Risk maps at Nevada also show that tornadoes with magnitude of EF2 and greater are rarely occur in this region, thus only EF0 and EF1 tornado induced wind should be considered for structures with a special requirement, for example, nuclear power plant.

The annual probability of experiencing EF3 and greater \((V \geq 136\text{ mph})\) tornadoes are presented in Figure 3.7 (D) to Figure 3.10 (D) for point target, small target, medium target and large target, respectively. The high risk region area is shrinking, while increasing the expected wind speed and the highest magnitude is \(10^{-4}\). The high risk regions are bounded by Rocky Mountains and Appalachian Mountains, and strong tornados (EF3 and
higher) are hardly visible in West Coast region except for the south California and north Arizona.

The annual probability of experiencing EF4 and greater \( (V \geq 166mph) \) tornadoes are presented in Figure 3.7 (E) to Figure 3.10 (E) for point target, small target, medium target and large target, respectively. These figures reveal that the highest striking probability still situated in Tornado Alley and Dixie Alley, however, the magnitude of hazard was drastically reduced to \( 10^{-5} \). Compare the result of EF0 to EF3 hazard maps, South California and north Arizona has much lower chance of experiencing EF4 or higher tornado because of the max rating tornado that observed in this region is EF3. The shrinkage of the high risk region also appears in Florida Peninsula, and most tornados in this region are induced by hurricane (Novlan & Gray, 1974). The observed tornados in north Wyoming have created a high hazard zone in these hazard maps, structure design in this zone should pay more attentions to tornado induced wind speed even though these regions are commonly considered excluded of tornado high risk regions.

The annual probability of experiencing EF5 and greater \( (V \geq 200mph) \) tornadoes are presented in Figure 3.7 (F) to Figure 3.10 (F) for point target, small target, medium target and large target, respectively. The probability of experiencing EF5 tornado in majority of United States is very low \( (10^{-10}) \), high risk region \( (10^{-6}) \) for point target include Kansas, Oklahoma, Iowa, southern Wisconsin, Tennessee, Mississippi, Alabama, and Georgia. The high risk region extends to the whole area of Tornado Alley and Dixie Alley for large target, plus Midwest and Kentucky. Overall, the coverage area and magnitude of EF5 tornado hazard is the smallest value, this happened because tornadoes
assigned an EF5 rating have historically been rare and only 59 tornadoes rated EF5/F5 since 1950.
Figure 3.7: Annual probability of experiencing an EF0-EF5 wind speed of point target in the continental United States.
Figure 3.8: Annual probability of experiencing an EF0-EF5 wind speed of small target in the continental United States.
Figure 3.9: Annual probability of experiencing an EF0-EF5 wind speed of medium target in the continental United States.
Figure 3.10: Annual probability of experiencing an EF0-EF5 wind speed of large target in the continental United States.
3.6.2 Tornado hazard estimation using hazard maps

Four locations are selected for comparing the tornado hazard of target with increasing size. Two location in high tornado risk region, one near Oklahoma City (OK), the other one near Birmingham (AL). One location in moderate tornado risk region, San Antonio (TX). The last one is in low tornado risk region, Bozeman (MT). Figure 3.11 (A) and Figure 3.11 (D) clearly show the striking probability of target with same size can be several orders of magnitude higher or lower in different location. According to Figure 3.11, tornado hazard increases with increasing target size, and the shape of the hazard curves are remained almost the same while size increasing. In Oklahoma City, the highest hazard is for the large target experiencing an EF0 or higher tornado, and the same hazard for the point target is about one order of magnitude decrease in probability level. The hazard curves in Oklahoma City and Birmingham are nearly identical. Tornado hazard in San Antonio is about one order of magnitude decrease compared with Oklahoma City. It has to note that although striking probability in Oklahoma City and in Birmingham are similar, the probability of having nocturnal tornados and off-season tornados is higher in Dixie Alley (Ashley et al., 2008).
3.6.3 Combined Hazard Curve

Tornado hazards are not currently required in the design wind load of building codes. To compare the simulated tornado hazard curve and the current suggested design wind speeds in building code (ASCE 7-16), results are shown in Figure 3.12 and Figure 3.13. The design wind speeds are provided in building code (ASCE) for different mean recurrence interval (MRI) (i.e. 10 years to 3000 years mean recurrence interval), and the corresponding annual probability of exceedance (POE) can be directly determined from Figure 3.11: Tornado wind speed hazard for different cities. Oklahoma City, OK(A); Birmingham, AL(B); San Antonio, TX(C); Bozeman, MT(D).
the MRI (POE = 1/MRI). The annual probability of non-exceedance is commonly modeled as a Type-I extreme value distribution (Gumbel distribution), which is given by:

\[ F(x) = \exp\{-\exp[-a(x - \mu)]\} \]

where \( x \) is the wind speed, \( a \) and \( \mu \) are the scale and location parameters that can be evaluated by the least-squares method. The design wind speed and the corresponding MRI in the US are obtained from the Basic Wind Speeds Map provided by ASCE 7-16 design code. The annual wind speed probability distributions for selected cities are fitted using Gumbel distribution, the required wind speed in building code (red dot) and the fitted distributions (red dash line) are plotted.

The annual probability of exceedance of tornado induced wind speeds are derived from the tornado hazard maps shown in section 3.6.1. Instead of directly fitting probability of exceedance, probability \( F(x) \) of experience a tornado with the peak wind speed less than a given value \( x \) can be fitted by exponentiated Weibull distribution which proposed by Mudholkar and Srivastava (1993). The distribution function is

\[ F(x) = [G(x)]^\nu = [1 - \exp(-(x/\beta)^\alpha)]^\nu \]

where \( x \) is the wind speed, \( G(x) \) is the standard two-parameter Weibull distribution, \( \alpha \) and \( \nu \) are the two shape parameters and \( \beta \) is the scale parameter. The annual probability of exceedance equals to \( 1 - F(x) \).

In high tornado risk region, such as Oklahoma City and Birmingham (Figure 3.12 A and B), tornado induced wind plays an important role for structures considered extreme event with MRI greater than 700 year or annual probability of occurrence as low as \( 1.4 \times 10^{-7} \).
(Risk Category II). For a point targets, tornadic wind speed governs when extreme events with a probability of occurrence equals to $1 \times 10^{-4}$/year are considered.

The hazard curves in Figure 3.12 also show that as the wind speed increasing, both tornado hazard curve and non-tornadic curve decreasing, however, the probability of exceedance of non-tornadic wind speeds decrease at a faster rate than tornado wind speeds. For example, the design basic wind speed from building code is greater than the tornado induced wind speed with a 25 year MRI in Oklahoma City (Figure 3.12 A) and the wind speed equals to 83 mph with an annual probability of exceedance of $4 \times 10^{-2}$ (ASCE 7-16), whereas for 0.62 mi$^2$ circular targets and point targets the tornado induced corresponding probability are $3 \times 10^{-3}$ and $4.2 \times 10^{-4}$ respectively. However, the tornado induced wind become more significant as the extreme wind speed increases. For instance, the design basic wind speed with a 3000 year MRI (Risk Category IV) in Oklahoma City is 121 mph with an annual probability of exceedance of $3.3 \times 10^{-4}$, whereas for 0.62 mi$^2$ circular targets and point targets the tornado induced probability are $9 \times 10^{-4}$ and $1.2 \times 10^{-4}$. This suggests that wind loading for very large structure with a high Risk Category could be dominated by tornadoes in high tornado risk region, such as Oklahoma City and Birmingham.

In moderate and low risk region, tornado induced wind speeds do not control the design wind load for common structures. The intersections of curves with non-tornadic wind and tornado induced wind are about $7 \times 10^{-5}$ for large target and $8 \times 10^{-7}$ for point target. In fact, the tornado induced wind load only need to take into consideration for special structures such as nuclear power plant where wind speed with $10^6$ and $10^7$ year MRI are required.
The design wind speeds at different locations in the United States are often developed based on wind speed data from certified stations. However, as mentioned previously, tornado induced wind speed is excluded because of the very low occurrence rate and the dimension of tornado which is too small to be captured by wind instruments. When the ignorance of tornado wind speed exists, structure engineers may underestimate the wind loads in designing buildings with high risk category in tornado-prone regions. The failure of these kinds of high risk buildings usually lead to significant economic losses or cause substantial hazard to the human life, such as chemical plant, hospital and nuclear power plant. Furthermore, the phenomenological basis of the tornado is very different from straight-line wind and reflects different damage cases based on the survey (Marshall, 2004). Thus, the complete distribution of extreme wind speeds at the building site is a mixed distribution which contains profile of the risk scenario from the non-tornado events and from the tornado events. The combined distribution of such dual extreme wind phenomena has been determined by the equation:

\[ P(V \leq v_i) = 1 - P(V \leq v_N) = 1 - P(V \leq v_{i_{Non-tor}}) \times P(V \leq v_{i_{Tor}}), \]

where \( P(V \leq v_i) \) is probability of wind speed less than the combined extreme wind speed, \( P(V \leq v_{i_{Non-tor}}) \) is nonexceedance probability of non-tornado wind, and \( P(V \leq v_{i_{Tor}}) \) is the nonexceedance probability of tornado wind. Figure 3.12 (B, D) and Figure 3.13 (B, D) shows the combined annual probability of exceedance of both type of wind in selected cities in United States.
Figure 3.12: Combined tornadoes and non-tornadic extreme wind speed distribution in Oklahoma City (A, B) and Birmingham (C, D).
The study investigated the spatial distribution of annual probability of exceedance of EF0 to EF5 tornadoes considering the structure size effect. Site- and structure size-specific tornado hazard maps were created using a simulated tornado database. It has been shown that the strike probability of tornadoes increases with increasing target size. Except for in Tornado Alley and Dixie Alley, weak tornadoes (EF0 and EF1) govern the tornado hazards of Western and Eastern coasts of the United States. The annual strike probability

![Combined tornadoes and non-tornadic extreme wind speed distribution in San Antonio (A, B) and Bozeman (C, D).](image)

**Figure 3.13:** Combined tornadoes and non-tornadic extreme wind speed distribution in San Antonio (A, B) and Bozeman (C, D).

### 3.7 Conclusion

The study investigated the spatial distribution of annual probability of exceedance of EF0 to EF5 tornadoes considering the structure size effect. Site- and structure size-specific tornado hazard maps were created using a simulated tornado database. It has been shown that the strike probability of tornadoes increases with increasing target size. Except for in Tornado Alley and Dixie Alley, weak tornadoes (EF0 and EF1) govern the tornado hazards of Western and Eastern coasts of the United States. The annual strike probability
of weak tornadoes is estimated to be in the order of $10^{-2}$ in high risk region. The high risk regions for strong tornadoes (EF2 and EF3) are bounded by Rocky Mountains and Appalachian Mountains, and the highest strike probabilities in these regions are about $10^{-3}$ to $10^{-4}$. Violent tornadoes (EF4 and EF5) are rare events which have a very low probability of occurrence in regions outside of Tornado Alley, Dixie Alley, and Midwest with the highest annual occurrence probability of approximately $10^{-5}$ to $10^{-6}$.

### 3.8 Reference


4.1 Abstract

United States has received a tremendous damage and loss in properties by tornado strikes for the past few decades. In order to predict the tornado damages and improve the community resilience performance, a new approach of tornado scenario selection and damage estimation is proposed in this study. Moore (OK) was selected as the study domain because of the high frequency of tornado occurrence, and a simulated tornado database which include 1 million years tornado tracks was applied to model the tornado risks in this region. The damage area and peak wind speed have been calculated, for each tornado tracks, to estimate its corresponding mean recurrence interval (MRI). The building locations and dimensions are determined using an image segmentation algorithm, and the damage state is evaluated using the fragility curves. Damage estimation for three tornado scenarios, selected according to damage area, peak wind speed and both intensity measures were conducted with different hazard level (MRI).

4.2 Introduction

Tornado causes significant damage to property and casualties every year in United State. Based on the damage report from National Oceanic and Atmospheric Administration (NOAA), there are about 1,200 tornadoes in the United States observed each year. As a
result, the average annual number of fatalities is more than 90 and injure approximately 1,500 as well as billions of dollars in economic losses (Simmons et al., 2010). A major reason of tornado strikes leads to such a severe consequences is the lack of awareness by public and of the quantitative insight into tornado damage assessment. For example, community could experience a huge tornado losses because of the location of critical buildings and the concentrated residential buildings. Therefore, tornado damage evaluation and prediction play a vital role in improving the community resilience performance and decreasing the economic losses or fatalities. Unfortunately, no study has been reported so far on tornado building damage prediction and mitigation, which is critical for emergency management planning and for insurance companies to estimate the potential payouts.

Tornado building damage prediction is an issue because of the difficulties in predicting the tornado tracks and in simulating the building location and dimension. In order to predict the potential tornado hazard in the target community, a simulated tornado database which contains 1 million years of tornado tracks is applied in this study. Owing to the greater ease of accurately extracting geographic data, image segmentation is introduced to determine the dimension and location of the buildings. For a target domain under tornado strike, the tornado damage area and peak wind speed are the two main parameters for determining the building damage, however, these two parameters are not highly correlated. In other words, tornado with a higher wind speed are not always result in a larger damage areas, and it is possible to see that an EF5 tornado only affect a small portion areas of study domain. For that reason, hazard consistent tornado scenarios (with similar mean recurrence interval (MRI)) are selected based on three intensity measures,
peak wind speed, damage area and joint occurrence of peak wind speed and damage area. In this study, the induced building damages for these three scenarios are compared with a different MRI from 10 years to 3000 years.

4.3 Methodology

The analytical procedure for quantitatively examine the tornado induced damage for the selected study domain has shown in Figure 4.1. The simulation includes two parts, tornado track selection and damage estimation, and the main steps are briefly discussed as follows:

(1) Define study domain (Section 4.4.1): Moore (OK) has selected as the study domain because of the frequent observed tornados in this area.

(2) Load tornado track database (Section 4.4.2): Tornado database with 1 million years simulation tracks is applied in this study.

(3) Determine tornado peak wind speed and impacted area (Section 4.4.3): A stationary wind field model (Modified Rankine Vortex Model) is used to determine the wind speed. Tornado footprint is idealized as a rectangle and the degradation of wind speed along the track is considered.

(4) Calculated Mean Recurrence Interval (MRI) (Section 4.4.5): Compute the site-specific MRI for each tornado tracks according to the peak wind speed and damage area obtained from Step 3.
(5) Select tornado tracks (Section 4.4.6): hazard consistent tornado scenarios (with similar MRI) are selected based on three intensity measures, peak wind speed, damage area and joint occurrence of peak wind speed and damage area.

(6) Determine building location and dimension within the study domain (Section 4.5.1): Building stocks are identified using the image segmentation method.

(7) Determine building fragility curves (Section 4.5.2): 14 types of building with 4 different damage states is considered.

(8) Damage estimation (Section 4.6.1): Tornado induced building damage is estimated according the peak wind speed occurred at the building sites and the fragility curves of the affected buildings.

Figure 4.1: Flowchart of tornado track selection and damage estimation procedures.
4.4  Tornado track selection

4.4.1 Study domain

The study domain is set to Moore, Oklahoma (Figure 4.2). The primary reason why Moore selected as the study domain is the geography and climate of the city. Moore is located in Central Oklahoma and at the south potion of Oklahoma City, where is in the center of what is colloquially known as Tornado Alley (Brooks et al., 2003; Gagan et al., 2010). The city has a population of approximately 55,000 and has a total area of 22.2 mi$^2$ according to the United States Census Bureau. The digitized study domain has shown in Figure 4.3, and the area is about 27.4 mi$^2$. The total area of the digitized domain is slightly larger than the area of city Moore, because the subdomains near the city are also included for simplicity.

Figure 4.2: Location of study domain.
4.4.2 **Tornado database**

The Storm Prediction Center (SPC), a division of the National Oceanic Atmospheric Administration (NOAA), maintains a tornado database of more than 60,000 tornado records since 1953. Unlike hurricanes and typhoons, tornado striking is short lived, covering and affecting small areas. Therefore, the historical database is still not taken over a long enough period to estimate the risk of strong tornadoes (EF4+) for tornado prone regions or weak tornado (EF1+) in tornado infrequent regions. The stochastic simulation framework to generate synthetic tornado tracks for the continental United States has been discussed in [CHAPTER 2](#), and the simulated tornado database of 1,000,000 years of simulated tornado tracks will applied in this study. Tornado occurrence rate was modeled using a negative binomial distribution. Probability density GIS maps, derived by using kernel density estimation (KDE), were applied in determine tornado spawn location and

Figure 4.3: Digitized study domain.
tornado parameters, such as dimension, time, intensity and direction. All these parameters are geographic dependent, in other words, the properties vary based on the geographic locations.

Simulated tornado tracks were selected if the impacted area overlaps with the study domain. A total of 106,401 tornado strike the target domain in 1,000,000 simulation years, which include 32,608 EF0 tornadoes, 45,011 EF1 tornadoes, 17,881 EF2 tornadoes, 7,457 EF3 tornadoes, 2,850 EF4 tornadoes, and 594 EF5 tornadoes (Figure 4.4). Figure 4.5 shows the accumulated tornado tracks of the 10,000 simulation years, and these tracks will be used for damage estimation.

![Figure 4.4: Distribution of tornado intensity in study domain.](image-url)
4.4.3 Calculate wind speed

Tornadoes do not usually maintain their peak intensity along the entire footprint, however only a small portion of area is impacted by the maximum intensity. In order to simulated the variation of tornado intensity and wind speed over its life cycle, the modified Rankine vortex model combined with the degradation model is applied in this study. Note that only tangential wind speed is considered in the modified Rankine vortex, and the swirl motion of the tornado can be reasonably approximate as:
\[ V_{\text{tan}}(r) = \frac{r \Gamma_\infty}{\pi (r^2 + r_c^2)} \]

where \( \Gamma_\infty = 2 \pi r_c V_{\text{tan,max}} \) is the circulation considered as a constant and \( r_c \) is the core radius where the maximum tangential velocity \( V_{\text{tan,max}} \) occurs. For a given tornado track, the only unknown in this equation is \( r_c \), and \( r_c \) can be determined by assuming the lower bound of EF0 occurs at the edge of the tornado path width.

The degradation model along the tornado track is adopted from study of Faletra et al., (2016) and study of Twisdale et al., (1981). In these studies, the wind speed variation along the length of tornado was determined and the detail have been shown in section 2.3.2.3.

4.4.4 Calculate impacted area

Besides the tornado peak wind speed, the size of the damage areas also plays an important role in tornado damage assessment. Tornado footprint is idealized as a rectangle defined by the tornado length \( L \) and tornado width \( W \) (Twisdale & Dunn, 1983). Figure 4.6 shows modeled tornado damage area, the solid blue line is the boundary of damage area and the dash line is the center line of path. When tornado strikes the study domain, only overlapping area is recorded (Figure 4.6); where peak wind speed is measured within the overlapping area. Therefore, the peak wind speed within the study domain could be lower than the peak wind speed of the tornado, and dependent on the striking location and direction.
From this point of view, the impact area and impact speed are two major factors in evaluated tornado induced damage. Large scale tornadoes cover larger areas under damage; while tornadoes with higher wind speed causes more damages for buildings inside of affected area.

The correlation coefficients ($\rho$) of tornado impact areas and peak wind speeds have shown in Table 4.1. The value for overall tornado is about 0.51 which can be considered moderately correlated, however, the value tends to decrease by involving the EF scale. In other words, tornado with a higher wind speed are not always result in a larger damage area, and it is possible to see that an EF5 tornado only affect a small portion of study domain. For this reason, both intensity measures, damage area and peak wind speed, should be considered for a given hazard level.

Figure 4.6: Modeled tornado impacted area.
4.4.5 Determine Mean Recurrence Interval

For each simulated tornado track, peak wind speed (4.4.3) and damage area (4.4.4) within the study domain was calculate to determine the site-specific Mean Recurrence Interval (MRI). The MRI is the expected time at which event of a given value or greater can occur, and the MRI of a selected tornado event \( i \) can be defined using the following equations:

\[
MRI(a_i > A) = \frac{1}{\lambda P(a_i > A)} = \frac{Y}{n_a}
\]  

(4.2)

\[
MRI(v_i > V) = \frac{1}{\lambda P(v_i > V)} = \frac{Y}{n_v}
\]  

(4.3)

\[
MRI(a_i > A \cap v_i > V) = \frac{1}{\lambda P(a_i > A \cap v_i > V)} = \frac{Y}{n_{va}}
\]  

(4.4)

where \( a_i \) and \( v_i \) are the tornado peak wind speed and damage area occurred in study domain for event \( i \). \( n_a \) is the number of tornadoes with a wind speed greater than \( A \). \( n_v \) is the number of tornadoes with a damage area greater than \( V \). \( n_{va} \) is number of tornadoes each having a peak wind speed greater than \( V \) and a damage area greater than \( A \).
Figure 4.7: determination of $n_\alpha, n_v$ and $n_{va}$ for a given boundary ($V, A$).

Figure 4.7 shows the determination of the number of tornadoes with a wind speed, damage area or both values greater a given data point. In Figure 4.7, tornado tracks with the corresponding peak wind speed and damage area within the study domain area plotted. For a given peak wind speed ($V$) and damage area ($A$), tornado data points located in the red region have peak wind speed greater than $V$ and damage area greater than $A$. Therefore, $n_{va}$ equals to the number of points in the red portion, while $n_\alpha$ equals to the number of points in the yellow portion plus the points in the red portion and $n_v$ equals to the number of points in the green portion plus the points in the red portion.

Figure 4.8 to Figure 4.10 present the MRI regarding to the tornado induced peak wind speed and damage area in Moore, Oklahoma. In these figures, the MRI for 300, 700,
1700 and 3000 years were plotted and the selected tornado candidates used in section 4.5 were picked at each of these MRI level.

Figure 4.8: MRI for tornado peak wind speed.

Figure 4.9: MRI for tornado damage area.
To perform the analysis of Moore subject to tornadoes, tornado scenarios were selected according to the peak wind speed, damage area, and both intensity measures for a given hazard level (MRI = 10, 25, 50, ..., 3000). Once the MRI values of each tornado event in the study domain were determined using Equation (4.2) to Equation (4.4), tornado events were then grouped according to the target MRI range and the number of events for each assemble is presented in Table 4.2.

Figure 4.10: Joint MRI for tornado peak wind speed and damage area.

4.4.6 Tornado tracks selection

To perform the analysis of Moore subject to tornadoes, tornado scenarios were selected according to the peak wind speed, damage area, and both intensity measures for a given hazard level (MRI = 10, 25, 50, ..., 3000). Once the MRI values of each tornado event in the study domain were determined using Equation (4.2) to Equation (4.4), tornado events were then grouped according to the target MRI range and the number of events for each assemble is presented in Table 4.2.
Table 4.2: Numbers of tornados scenarios selected for a given MRI range.

<table>
<thead>
<tr>
<th>MRI range (year)</th>
<th>Number of tornado scenarios</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRI($a_i &gt; A$)</td>
<td>MRI($v_i &gt; V$)</td>
</tr>
<tr>
<td>10 ± 1</td>
<td>15,489</td>
</tr>
<tr>
<td>25 ± 2.5</td>
<td>8,082</td>
</tr>
<tr>
<td>50 ± 5</td>
<td>4,041</td>
</tr>
<tr>
<td>100 ± 10</td>
<td>2,021</td>
</tr>
<tr>
<td>300 ± 30</td>
<td>673</td>
</tr>
<tr>
<td>700 ± 70</td>
<td>289</td>
</tr>
<tr>
<td>1700 ± 170</td>
<td>119</td>
</tr>
<tr>
<td>3000 ± 300</td>
<td>67</td>
</tr>
</tbody>
</table>

4.5 Damage estimation

4.5.1 Building location

In order to perform the site-specific building damage assessment, building locations were identified by using the image segmentation toolbox in MATLAB. Image segmentation is the process of dividing a digital image into multiple parts of regions. The primary goal of using this method is to locate the building and its boundary. Thresholding method is applied to binarize the map, pixel values in the image that are less than or equal to the threshold value are replaced with black pixel, or a white pixel if the pixel values are greater than that threshold value. Once the binary version of the original image is obtained, the boundaries and locations of these objects can be identified. The main procedures of image segmentation are presented as follows and illustrated in Figure 4.11:

1. Locate the study domain and remove the landmarks and labels using the Google Maps APIs Styling Wizard (Figure 4.11 (A) and (B)).
(2) Load the map into MATLAB and convert the image into a binary image using the thresholding method (Figure 4.11 (C)).

(4) Identify buildings and save the building boundaries in the binary image.

(3) Mapping the building boundaries back to its original coordinate using the map scale (Figure 4.11 (D)).

Figure 4.11: Image identification and segmentation using MATLAB.
Figure 4.12 shows a remarkable agreement between the building locations and the identified building boundaries using image segmentation method. The total number of identified buildings in Moore is 22,753, which is close to the reported number of 21,444 from United States Census Bureau.

Figure 4.12: Identified building boundaries using image identification method.

4.5.2 Building type and fragility curves

In order to accurately estimate the tornado damage, it is necessary to generate a site-specific study domain especially for the location of critical buildings such as schools, hospitals, and fire station. The modeled building types applied in this study have 14 categories, which include residential building, small commercial building, light industry building, elementary school, high school, fire station, hospital, community center/church,
government office, and shopping center. The number of building stocks for each category has present in Table 4.3, and the locations have been plot in Figure 4.13. Note that the building type 6 to type 14 have been manually assigned to the identified locations inside of study domain according to the corresponding geographic locations, and type 1 to type 5 are randomly assigned to the remaining locations.

Figure 4.13: Buildings in Moore, Oklahoma.
Each building contains 4 damage levels, slight, moderate, extensive, and complete. Slight damage (Damage State 1) means minor damage at doors and windows/roof occurred but repairable and can be re-occupied immediately. Moderate damage (Damage State 2) represents for the moderate damage on window, door and roof covering occurred but repairable and able to be occupied. Extensive damage (Damage State 3) describes the severe damage on building envelops and cannot be unless repaired. Complete Damage (Damage State 4) means the buildings completely leveled and cannot be repaired. The adapted fragility curves were developed to quantitatively describing these levels of damage according to the performance of building envelope, roof and walls.

Table 4.3: Building Types and Description. (Adapted from Memari et al 2018.)

<table>
<thead>
<tr>
<th>Building Type</th>
<th>Building Description</th>
<th>Number of units</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Residential, wood, 1 story, small size, gable roof</td>
<td>4420</td>
</tr>
<tr>
<td>2</td>
<td>Residential, wood, 2 stories, small size, gable roof</td>
<td>4446</td>
</tr>
<tr>
<td>3</td>
<td>Residential, wood, 1 story, medium size, gable roof</td>
<td>4389</td>
</tr>
<tr>
<td>4</td>
<td>Residential, wood, 2 stories, medium size, hip roof</td>
<td>4455</td>
</tr>
<tr>
<td>5</td>
<td>Residential, wood, 2 stories, large size, gable roof</td>
<td>4404</td>
</tr>
<tr>
<td>6</td>
<td>Small commercial building</td>
<td>352</td>
</tr>
<tr>
<td>7</td>
<td>Light Industry building</td>
<td>189</td>
</tr>
<tr>
<td>8</td>
<td>Elementary school</td>
<td>10</td>
</tr>
<tr>
<td>9</td>
<td>High school</td>
<td>4</td>
</tr>
<tr>
<td>10</td>
<td>Fire station</td>
<td>3</td>
</tr>
<tr>
<td>11</td>
<td>Hospital</td>
<td>3</td>
</tr>
<tr>
<td>12</td>
<td>Community center/church</td>
<td>20</td>
</tr>
<tr>
<td>13</td>
<td>Government office</td>
<td>3</td>
</tr>
<tr>
<td>14</td>
<td>Shopping center</td>
<td>55</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>22753</strong></td>
</tr>
</tbody>
</table>

Fragility curves are commonly used in risk assessment for evaluating the performance of buildings, the curve represents the probability of exceeding a given damage
state due to various wind speed. Memari et al. (2018) have summarized and developed tornado fragility curves for 19 types of building according to the previous study (Amini et al., 2014; Koliou et al., 2017; Masoomi et al., 2016).

Fragility curves for structure system or component are typically expressed analytically by a lognormal cumulative distribution function (CDF) (e.g., Roueche et al. 2017; Amini et al. 2014; Ellingwood et al. 2004), which is described as follows:

\[ F(x) = \Phi \left[ \frac{\ln(x) - \mu}{\sigma} \right] \]  \hspace{1cm} (4.5)

where \( x \) is the 3-s gust wind speed (m/s or mph) for tornado; \( \Phi(.) \) is the standard normal CDF; \( \mu \) is logarithmic median of structure capacity; \( \sigma \) is logarithmic standard deviation of structure capacity.

Because the real dimension of the building is considered in this study, it is possible to see that a building experience different wind speeds during a tornado impact. The simplification inherent in this calculation is to only consider the maximum wind speed for the impact building. To obtain the peak wind speed, the target building site is further mesh to 10m x 10m grids (see Figure 4.14). Wind speed at each grid point are evaluated and only the maximum wind speed is recorded for the damage estimation.
Figure 4.15 presents a simulated EF5 tornado occurred at the study domain and the induced building damages. Buildings with a high damage state rate are often observed at the tornado core area, and the probability of having a severely damaged building decreases with the decreasing wind speed.
4.6 Result and Discussion

The EF scale is the most common way to measure the tornado magnitude based on wind damage. To perform the analysis of building stocks in city Moore subject to tornadoes, simulated tornado tracks are grouped according to its peak EF scale. It has to note that, the peak EF scale is the maximum rating of the whole footprint and the area experiencing the peak EF scale could occurs outside of the study domain. Also, simulated tornado which does not cause any building damage are excluded in this part of study, and the number of the rest of the tornado tracks is shown in Table 4.4. Table 4.4 also shows the percentage of tornados which cause building damages, weak tornadoes have a low percentage because of
the smaller coverage area and induced wind speed compared to the high magnitude tornados.

Table 4.4: Numbers of tornadoes induced building damages.

<table>
<thead>
<tr>
<th>EF scale</th>
<th>EF0</th>
<th>EF1</th>
<th>EF2</th>
<th>EF3</th>
<th>EF4</th>
<th>EF5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Counts</td>
<td>20,878</td>
<td>35,678</td>
<td>15,308</td>
<td>6,809</td>
<td>2,629</td>
<td>559</td>
</tr>
<tr>
<td>Percentage</td>
<td>64%</td>
<td>79%</td>
<td>86%</td>
<td>91%</td>
<td>92%</td>
<td>94%</td>
</tr>
</tbody>
</table>

Table 4.5 to Table 4.10 summarize the effect of tornadoes on the different types of building, showing the average number of buildings in each damage state (DS) induced by tornadoes from EF0 to EF5. For weak tornadoes (EF0 and EF1), the majority of the damaged buildings are classified as damage state 1 (DS1), DS2 and higher level damage states are barely observed for all types of buildings except residential buildings, and none of the simulated tornadoes in this category cause DS3 or DS4 for hospitals. With regard to strong tornadoes (EF2 and EF3), the main damage types are shifted to DS2 and higher, although the numbers of buildings in DS1 are still high. For violent tornado (EF4 and EF5), the average numbers of damaged buildings are keep increasing, DS1 to DS4 have been observed at all types of buildings and the numbers of buildings experiencing DS4 are increased dramatically.

Figure 4.16, from another point of view, presents the relationship between building damage states and the tornado magnitude. EF2 and lower tornadoes account for more than 50% buildings in DS1. EF2 and lower tornadoes still account for the majority of DS2 except for residential buildings, hospitals and fire stations. DS3 is dominated by EF3 and greater tornadoes for residential buildings, small commercial buildings, fire stations and
hospitals; about 60% of damaged buildings, for remaining types, are induced by EF3 and greater tornadoes. More than 50% of residential buildings, high schools and hospital experiencing DS4 as a result of hit by EF4 and greater tornadoes, and the remaining types of building are controlled by EF3 and greater tornadoes.

Figure 4.16: Tornado caused building damages grouped by different Damage State (DS).
Table 4.5: Average numbers of damaged buildings caused by EF0 tornado.

<table>
<thead>
<tr>
<th>Total</th>
<th>1-5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>DS1</td>
<td>6.89</td>
<td>5.81</td>
<td>0.583</td>
<td>0.224</td>
<td>0.0307</td>
<td>0.0138</td>
<td>3.21e-3</td>
<td>2.78e-3</td>
<td>0.0638</td>
<td>6.61e-3</td>
</tr>
<tr>
<td>DS2</td>
<td>1.01</td>
<td>0.700</td>
<td>0.201</td>
<td>0.0272</td>
<td>1.92e-3</td>
<td>1.72e-3</td>
<td>0</td>
<td>1.44e-4</td>
<td>0.0105</td>
<td>3.98e-3</td>
</tr>
<tr>
<td>DS3</td>
<td>0.0224</td>
<td>0.0172</td>
<td>0</td>
<td>4.79e-3</td>
<td>0</td>
<td>9.58e-5</td>
<td>0</td>
<td>1.44e-4</td>
<td>4.79e-5</td>
<td>4.79e-5</td>
</tr>
<tr>
<td>DS4</td>
<td>3.98e-3</td>
<td>8.62e-4</td>
<td>2.01e-3</td>
<td>5.75e-4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5.27 e-4</td>
<td></td>
</tr>
<tr>
<td>All DS</td>
<td>7.93</td>
<td>6.53</td>
<td>0.786</td>
<td>0.257</td>
<td>0.0327</td>
<td>0.0157</td>
<td>0.00321</td>
<td>0.00292</td>
<td>0.0745</td>
<td>0.0106</td>
</tr>
</tbody>
</table>

Table 4.6: Average numbers of damaged buildings caused by EF1 tornado.

<table>
<thead>
<tr>
<th>Total</th>
<th>1-5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>DS1</td>
<td>25.8</td>
<td>23.6</td>
<td>1.18</td>
<td>0.566</td>
<td>0.0590</td>
<td>0.0237</td>
<td>8.83e-3</td>
<td>7.46e-3</td>
<td>0.124</td>
<td>0.0140</td>
</tr>
<tr>
<td>DS2</td>
<td>7.82</td>
<td>6.63</td>
<td>0.784</td>
<td>0.138</td>
<td>0.0113</td>
<td>8.21e-3</td>
<td>4.76e-4</td>
<td>2.10e-3</td>
<td>0.0419</td>
<td>0.0121</td>
</tr>
<tr>
<td>DS3</td>
<td>0.927</td>
<td>0.827</td>
<td>0.019</td>
<td>0.0603</td>
<td>2.89e-3</td>
<td>1.26e-3</td>
<td>0</td>
<td>4.65e-3</td>
<td>1.60e-3</td>
<td>9.75e-3</td>
</tr>
<tr>
<td>DS4</td>
<td>0.144</td>
<td>0.0746</td>
<td>0.0323</td>
<td>0.0308</td>
<td>3.08e-4</td>
<td>0</td>
<td>1.12e-4</td>
<td>0</td>
<td>4.48e-4</td>
<td>1.40e-4</td>
</tr>
<tr>
<td>All DS</td>
<td>34.7</td>
<td>31.1</td>
<td>2.01</td>
<td>0.796</td>
<td>0.0735</td>
<td>0.0332</td>
<td>9.42e-3</td>
<td>9.56e-3</td>
<td>0.171</td>
<td>0.0278</td>
</tr>
</tbody>
</table>

Table 4.7: Average numbers of damaged buildings caused by EF2 tornado.

<table>
<thead>
<tr>
<th>Total</th>
<th>1-5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>DS1</td>
<td>67.3</td>
<td>63.8</td>
<td>1.92</td>
<td>0.947</td>
<td>0.0806</td>
<td>0.0308</td>
<td>0.0198</td>
<td>0.0116</td>
<td>0.212</td>
<td>0.0241</td>
</tr>
<tr>
<td>DS2</td>
<td>47.0</td>
<td>44.1</td>
<td>2.00</td>
<td>0.337</td>
<td>0.0337</td>
<td>0.0219</td>
<td>6.01e-3</td>
<td>0.0174</td>
<td>0.101</td>
<td>0.0170</td>
</tr>
<tr>
<td>DS3</td>
<td>26.5</td>
<td>25.4</td>
<td>0.519</td>
<td>0.280</td>
<td>0.0330</td>
<td>0.0203</td>
<td>8.49e-4</td>
<td>0.000131</td>
<td>0.0592</td>
<td>0.0118</td>
</tr>
<tr>
<td>DS4</td>
<td>4.98</td>
<td>3.93</td>
<td>0.395</td>
<td>0.552</td>
<td>9.21e-3</td>
<td>2.61e-4</td>
<td>1.05e-3</td>
<td>0</td>
<td>9.21e-3</td>
<td>2.81e-3</td>
</tr>
<tr>
<td>All DS</td>
<td>146</td>
<td>137</td>
<td>4.84</td>
<td>2.12</td>
<td>0.157</td>
<td>0.0733</td>
<td>0.0277</td>
<td>0.0292</td>
<td>0.381</td>
<td>0.0557</td>
</tr>
</tbody>
</table>
Table 4.8: Average numbers of damaged buildings caused by EF3 tornado.

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>1-5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>DS1</td>
<td>170</td>
<td>162</td>
<td>4.28</td>
<td>2.22</td>
<td>0.156</td>
<td>0.0510</td>
<td>0.0515</td>
<td>0.0279</td>
<td>0.477</td>
<td>0.0511</td>
<td>0.644</td>
</tr>
<tr>
<td>DS2</td>
<td>127</td>
<td>120</td>
<td>4.72</td>
<td>0.888</td>
<td>0.0627</td>
<td>0.0388</td>
<td>0.0191</td>
<td>0.0426</td>
<td>0.227</td>
<td>0.0430</td>
<td>0.745</td>
</tr>
<tr>
<td>DS3</td>
<td>145</td>
<td>142</td>
<td>2.09</td>
<td>0.738</td>
<td>0.100</td>
<td>0.0703</td>
<td>0.00837</td>
<td>0.0101</td>
<td>0.203</td>
<td>0.0254</td>
<td>0.405</td>
</tr>
<tr>
<td>DS4</td>
<td>65.3</td>
<td>59.5</td>
<td>2.44</td>
<td>2.648</td>
<td>0.0599</td>
<td>0.00676</td>
<td>0.0107</td>
<td>0.000294</td>
<td>0.090</td>
<td>0.0238</td>
<td>0.487</td>
</tr>
<tr>
<td>All DS</td>
<td>507.3</td>
<td>484</td>
<td>13.53</td>
<td>6.50</td>
<td>0.379</td>
<td>0.167</td>
<td>0.0897</td>
<td>0.0809</td>
<td>0.997</td>
<td>0.143</td>
<td>2.28</td>
</tr>
</tbody>
</table>

Table 4.9: Average numbers of damaged buildings caused by EF4 tornado.

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>1-5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>DS1</td>
<td>244.4</td>
<td>231.5</td>
<td>7.06</td>
<td>3.49</td>
<td>0.229</td>
<td>0.0757</td>
<td>0.0711</td>
<td>0.0304</td>
<td>0.808</td>
<td>0.0936</td>
<td>1.07</td>
</tr>
<tr>
<td>DS2</td>
<td>176.1</td>
<td>166.9</td>
<td>6.41</td>
<td>1.22</td>
<td>0.0829</td>
<td>0.0468</td>
<td>0.0243</td>
<td>0.0673</td>
<td>0.315</td>
<td>0.0692</td>
<td>0.936</td>
</tr>
<tr>
<td>DS3</td>
<td>199.9</td>
<td>195.1</td>
<td>2.75</td>
<td>0.973</td>
<td>0.134</td>
<td>0.0822</td>
<td>0.0129</td>
<td>0.0183</td>
<td>0.252</td>
<td>0.0373</td>
<td>0.507</td>
</tr>
<tr>
<td>DS4</td>
<td>156.1</td>
<td>146.4</td>
<td>4.28</td>
<td>4.25</td>
<td>0.120</td>
<td>0.0247</td>
<td>0.0213</td>
<td>0.00609</td>
<td>0.186</td>
<td>0.0384</td>
<td>0.787</td>
</tr>
<tr>
<td>All DS</td>
<td>776.5</td>
<td>740</td>
<td>20.5</td>
<td>9.92</td>
<td>0.565</td>
<td>0.230</td>
<td>0.130</td>
<td>0.122</td>
<td>0.156</td>
<td>0.239</td>
<td>3.30</td>
</tr>
</tbody>
</table>

Table 4.10: Average numbers of damaged buildings caused by EF5 tornado.

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>1-5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>DS1</td>
<td>356.2</td>
<td>335.7</td>
<td>11.1</td>
<td>5.70</td>
<td>0.383</td>
<td>0.122</td>
<td>0.120</td>
<td>0.0322</td>
<td>1.175</td>
<td>0.136</td>
<td>1.59</td>
</tr>
<tr>
<td>DS2</td>
<td>236.0</td>
<td>223.1</td>
<td>9.20</td>
<td>1.51</td>
<td>0.113</td>
<td>0.0680</td>
<td>0.0215</td>
<td>0.0930</td>
<td>0.438</td>
<td>0.0966</td>
<td>1.40</td>
</tr>
<tr>
<td>DS3</td>
<td>286.5</td>
<td>280.3</td>
<td>3.70</td>
<td>1.04</td>
<td>0.165</td>
<td>0.106</td>
<td>0.0233</td>
<td>0.0376</td>
<td>0.340</td>
<td>0.0465</td>
<td>0.658</td>
</tr>
<tr>
<td>DS4</td>
<td>340.0</td>
<td>324.1</td>
<td>7.46</td>
<td>6.32</td>
<td>0.197</td>
<td>0.0519</td>
<td>0.0286</td>
<td>0.0268</td>
<td>0.333</td>
<td>0.0769</td>
<td>1.33</td>
</tr>
<tr>
<td>All DS</td>
<td>1218.6</td>
<td>1163.3</td>
<td>31.5</td>
<td>14.6</td>
<td>0.857</td>
<td>0.347</td>
<td>0.192</td>
<td>0.190</td>
<td>2.29</td>
<td>0.356</td>
<td>4.99</td>
</tr>
</tbody>
</table>
4.6.1 Damage estimation based on Mean Recurrence Interval

Tornado tracks have been sorted by the Mean Recurrence Interval (MRI) in three different intensity measures, namely “Damage Area”, “Wind Speed” and “Joint (Damage Area and Wind Speed)”. For each intensity measure, tornadoes are selected by the target MRI according to the method discussed in section 4.4.5. There are totally 8 target MRIs applied in this study, and building damages have been evaluated at each group of tornado tracks, with the same MRI, separately. The result have been summarized in Figure 4.17 and Figure 4.18, which showing the average number of damaged buildings evaluate for each intensity measure with different MRIs. These figures show that number of damaged building induced from scenario “Wind Speed” is higher than the other two cases when the MRI is less than or equal to 25 years. The average number of damaged buildings tend to increase with increasing MRI for all three scenarios, however the curves derived from “Damage Area” increase faster than the other two scenarios and dominate the average number of damage buildings when MRI is greater than 700 years. Similar trends could be observed in Figure 4.19, which present the average total number of damaged buildings per track for each scenario. For example, in damage state 1, when MRI equals 25 year, the average number of total damaged building per tornado track equals 13.6, 30.6 and 6.2 for scenarios of “Damage Area”, “Wind Speed” and “Combined” respectively; when MRI equals 700 year, the values are 455.9, 200, 231 for scenarios of “Damage Area”, “Wind Speed” and “Combined” respectively.
Figure 4.17: Average numbers of damaged buildings caused by three scenarios
Figure 4.18: Average numbers of damaged buildings caused by three scenarios
In this study, a site-specific tornado damage assessment was performed in Moore Oklahoma. In order to accurately simulate the tornado damage, building locations and dimensions are determined according to the corresponding geographic location using the image segmentation method. Therefore, size effect which mentioned in section 3.4.3 is also considered for all buildings in this study. There are 14 types of building assigned in the study domain which includes residential building, small commercial building, light
industry building, elementary school, high school, fire station, hospital, community center/church, government office, and shopping center. The damage state of each building contains four different level, namely slight, moderate, extensive, and complete damage. The fragility curves, which describe the damage states, were applied to quantitatively assess the building damages in the study domain. The hazard consistent tornado scenarios (i.e. with almost the same MRI) are consider according to the “Wind speed”, “Damage area” and joint occurrence probability of both intensity measures. It has been shown that “Wind speed” scenarios tend to cause more damages for tornado events with a MRI less than 50 years, while the “Damage area” scenarios control the building damages for tornado events with a MRI greater than 3000 years.

The methodology present in this study can be used to estimate any other areas for tornado regional damage assessment. The result of this study could be further applied to evaluate the losses, fatalities, and recovery of communities for a given tornado hazard level (MRI or EF scale).

4.8 Reference


CHAPTER 5
CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions and summary

In order to examine the tornado risks in United States, a tornado simulation program was developed and the simulated tornado database was further used to generate the hazard maps and to predict the property damages for tornado prone region. In the tornado simulation program, the annual spawn rate was sampled from a negative binomial distribution. The spawn or touchdown locations were simulated using geographic dependent kernel density estimation (KDE) GIS maps, which specifically account for the variability of tornado properties at different geographic regions (e.g. tornado alley, Dixie alley and etc.). The Rankin Vertex model was applied to determine the tornado wind speed, and a degradation of wind speed along the track was modeled using the study from Twisdale (1983). Finally, a 1 million years tornado database was derived and the simulated tornado information includes tornado occurrence rate, intensity (EF scale), touchdown location, touchdown time and path direction. All these parameters are geographic dependent, meaning the properties vary depend on the geographic locations. Good agreements were found between the simulated tornado database and the observed tornado spawn rate and other parameters.

To evaluate the potential tornado risks in United States, a series of tornado hazards was carried out using the simulated tornado database, and the hazard maps contain features that could accurately reflect the site- and size- specific wind probabilities. By examine the existing structures, four target building size was applied to cover the common structure
dimensions, namely, point target, small target (0.0096 mi²), medium target (0.038 mi²), and large target (0.5 mi²). For each target, the annual probability of exceeding a given wind speed was determined, and the overall probability is increased by increasing the target size and decreased by increasing the wind speed. According to the hazard maps, hazards of weak tornados (EF0 and EF1) covers almost all areas of United States and the highest magnitude of striking probability is up to 10⁻² in high risk region. The high risk regions for strong tornadoes (EF2 and EF3) are bounded by Rocky Mountains and Appalachian Mountains, and the highest magnitude of striking probability is about 10⁻³ to 10⁻⁴ according to the EF scale. Violent tornadoes (EF4 and EF5) are rare events which has a very low probability of occurrence in regions outside of Tornado Alley, Dixie Alley, and Midwest; the highest magnitude of the annual hazard is approximately 10⁻⁵ to 10⁻⁶. Using those hazard maps, tornado hazard curves were combined with the current suggested design wind in building code. The combined hazard curves indicated that tornado wind speed plays an important role for structures considered extreme event with MRI greater than 700 year or annual probability of occurrence as low as 1.4e⁻³ (Risk Category II) in high risk region such as Oklahoma City and Birmingham; while tornado wind speed controls the wind load in low tornado risk region only if the critical building, such as Nuclear power plant is designed.

One of the principle applications of the simulated tornado database is in damage prediction and assessment. To perform such a damage determination, building location and dimension are determined according to the corresponding geographic location using the image segmentation method. Different building types were assign to the modeled buildings and the potential damage was simulated using 4 fragilities curves, which describe
the damage states from slight to severe. During a tornado strike, the tornado damage area and peak wind speed are the two main parameters for determining the building damage, however, these two parameters are not highly correlated. In other words, tornado with a higher wind speed are not always result in a larger damage areas, and it is possible to see that an EF5 tornado only affect a small portion areas of study domain. For that reason, hazard consistent tornado scenarios (with similar mean recurrence interval (MRI)) are selected based on three intensity measures, peak wind speed, damage area and joint occurrence of peak wind speed and damage area. In this study, the induced building damage for these three scenarios are compared with a different MRI from 10 years to 3000 years. The tornado property damage assessment shows that “Wind speed” scenarios tend to cause more damages for tornado events with a MRI less than 50 years, while the “Damage area” scenarios control the building damages for tornado events with a MRI greater than 3000 years.

5.2 Recommendations for the future works

To improve and expand the current work discussed in this study, the following research topics are recommended for further work:

(1) It has to note that the stochastic tornado simulation model in this study is based on historical tornado reporting or observation. Though the SPC tornado database contains more than 60 years of observation data, the length of the database is still not long enough to accurately describe the characteristics of the rare tornado event (EF4+ tornado), and to account for any climatological change effects. Also, the EF scale rating system is purely
based on the post-event damage surveys, and there are many tornadoes could not be rated because lack of damage indicators. Therefore, the proposed model could be improved incorporating the data from real-time wind speed measure system (e.g. Doppler on Wheels network) and increasing the observed tornado record time.

(2) The idealized tornado footprint is a rectangle which defined by the tornado length and width. However, this is not always a good assumption to describe the tornado coverage area, especially when the tornado heading direction, wind speed, and dimensions are varying during its life-cycle. Future works should simulate the variation at each time step as a stochastic or random process to capture the complexity of real tornado footprint.

(3) In addition to producing uniform hazard design wind speed maps, the development of uniform risk design wind speed maps could be important to achieving a consistent performance of building throughout a nation under tornado wind load, even if the building located in regions with different tornado hazard. Through combining the tornado hazard curves (derived from hazard maps) at the building site with the corresponding structure fragility curves, the risk of tornado induced damage for the specific building can be evaluated and managed.
APPENDIX A
The density contours of tornado spawn location grouped by different heading angle

Wind Direction From 0° to 22.5°

Wind Direction From 22.5° to 45°
Wind Direction From 45° to 67.5°

Wind Direction From 67.5° to 90°
APPENDIX B

The density contours of tornado spawn location grouped by spawn month
APPENDIX C

The density contours of tornado spawn location grouped by spawn time
APPENDIX D
The density contours of tornado spawn location grouped by EF scale